THE RUELLE OPERATOR, ZETA FUNCTIONS AND THE ASYMPTOTIC DISTRIBUTION OF CLOSED ORBITS

Mark Pollicott

Submitted for the degree of Ph.D.
March 1981

Department of Mathematics,
University of Warwick,
Coventry, ENGLAND.

CONTENTS

Acknowledgements	(i)
Declaration	(ii)
Summary	(iii)
Chapter 1 A complex Ruelle-Perron-Frobenius theorem and two counter-examples	
0. Introduction	1.1
1. The Ruelle operator theorem	1.3
2. A complex Ruelle operator theorem	1.8
3. Extending the zeta function	1.15
4. A counter-example to Ruelle's question	1.20
5. A counter-example to Bowen's question	1.23
References	1.27
Chapter 2 An analogue of the prime number theorem for closed orbits of Axiom A flows.	
0. Introduction	2.1
1. Axiom A flows	2.2
2. Shifts of finite type	2.3
3. Axiom A flows and suspensions	2.4
4. The zeta function	2.5
5. Locally constant functions	2.6
6. When ϕ is not topologically weak-mixing	2.8
7. Ruelle's theorem	2.9
8. Extensions beyond $\Re(s) = h$	2.11
9. Extending ζ for an Axiom A flow	2.15
10. Number theory	2.15
References	2.19

Chapter 3 Asymptotic distribution of closed geodesics

0.Introduction 3.1

1.Closed orbits for suspended flows 3.3

2.Symbolic dynamics for geodesic flows 3.9

3.Closed orbits and zeta functions 3.14

4.Asymptotic results for closed geodesics 3.19

5.Further distribution results 3.21

References 3.27

Appendix A survey of symbolic dynamics for Axiom A flows

0.Reflections on Axiom A flows A.1

1.Axiom A flows

1.1 Definitions A.3

1.2 Canonical co-ordinates A.4

2.Two important properties of Axiom A flows

2.1 Expansiveness A.6

2.2 The tracing property A.6

3.Differential topology (on the manifold) A.8

4.Markov partitions (for the basic set)

4.1 Maps between sections A.11

4.2 Construction of chains A.13

4.3 The Markov patition A.15

5. Symbolic dynamics
5.1 Constructing the shift space A.19
5.2 Invariance of boundaries A.22
5.3 The semi-conjugacy map A.22
5.4 The semi-conjugacy map is bounded-one A.23
5.5 Asymptotic comparisons of closed orbits A.24
6. Manning's lemma
6.1 Preimages of the semiconjugacy A.26
6.2 Construction of a new shift space A.28
6.3 Construction of other shift spaces A.30
6.4 The counting lemma A.32
References A.34

5.Symbolic dynamics

5.1 Constructing the shift space A.19

5.2 Invariance of boundaries A.22

5.3 The semi-conjugacy map A.22

5.4 The semi-conjugacy map is bounded-one A.23

5.5 Asymptotic comparisons of closed orbits A.24

6.Manning's lemma

6.1 Preimages of the semiconjugacy A.26

6.2 Construction of a new shift space A.28

6.3 Construction of other shift spaces A.30

6.4 The counting lemma A.32

References A.34

Acknowledgements

I would like to thank the S.E.R.C. for their financial support.

I would like to express my gratitude to Prof. William Parry for his very considerable encouragement, interest and help. I consider myself fortunate to have him as a supervisor, collaborator and friend.

DECLARATION

The paper constituting Chapter One has been accepted for publication in Ergodic Theory and Dynamical Systems.

The paper which forms Chapter Two is to appear in Annals of Mathematics. This paper resulted from a collaboration with William Parry. Each author worked separately on different parts of the paper. Some of the material in sections 5 and 10 has appeared in a paper by Bill (Israel Journal of Mathematics, 45 (1983) 41-52.).

Chapter Three has yet to be submitted.

Some of the material in this thesis formed part of an M.Sc. course delivered by the author at Warwick University (January - March 1984).

SUMMARY

This thesis is composed of three independent chapters and an appendix. Each chapter has its own introduction, references and notation.

In chapter One a new proof of a theorem of Ruelle about real Perron-Frobenius type operators is given. This theorem is then extended to complex Perron-Frobenius type operators in analogy with Wielandt's theorem for matrices. Finally two questions raised by Ruelle and Bowen concerning analyticity properties of zeta functions for flows are answered.

In Chapter Two we improve a result of Ruelle on the domain of analyticity of the zeta function for an Axiom A flow. The method used requires results on complex Perron-Frobenius operators derived in the first chapter. These results are reproduced with alternative proofs. Finally, asymptotic estimates for numbers of closed orbits are deduced by analogy with the prime number theorem. This extends a result of Margulis.

In the first section of Chapter Three we give a relationship between periodic points and certain equilibrium states for subshifts of finite type. We next study geodesic flows on surfaces of constant negative curvature. We compare the zeta functions of a geodesic flow and a certain suspension flow. These results are then used to recover asymptotic estimates by Margulis and Bowen on the distribution of closed geodesics. Finally new results are given in two special cases.

The Appendix is an outline of Bowen's symbolic dynamics for Axiom A flows. This material is purely expository

SUMMARY

This thesis is composed of three independent chapters and an appendix. Each chapter has its own introduction, references and notation.

In chapter One a new proof of a theorem of Ruelle about real Perron-Frobenius type operators is given. This theorem is then extended to complex Perron-Frobenius type operators in analogy with Wielandt's theorem for matrices. Finally two questions raised by Ruelle and Bowen concerning analyticity properties of zeta functions for flows are answered.

In Chapter Two we improve a result of Ruelle on the domain of analyticity of the zeta function for an Axiom A flow. The method used requires results on complex Perron-Frobenius operators derived in the first chapter. These results are reproduced with alternative proofs. Finally, asymptotic estimates for numbers of closed orbits are deduced by analogy with the prime number theorem. This extends a result of Margulis.

In the first section of Chapter Three we give a relationship between periodic points and certain equilibrium states for subshifts of finite type. We next study geodesic flows on surfaces of constant negative curvature. We compare the zeta functions of a geodesic flow and a certain suspension flow. These results are then used to recover asymptotic estimates by Margulis and Bowen on the distribution of closed geodesics. Finally new results are given in two special cases.

The Appendix is an outline of Bowen's symbolic dynamics for Axiom A flows. This material is purely expository

When I have fears that I may cease to be
Before my pen has glean'd my teeming brain,
Before high-piled books, in charactery,
Hold like rich garners the full ripen'd grain

. . . . .

\- then on the shore
Of the wide world I stand alone, and think
Till love and fame to nothingness do sink

John Keats

A COMPLEX RUELLE-PERRON-FROBENIUS THEOREM AND TWO COUNTEREXAMPLES

Mark Pollicott

The operator $\mathcal{L}_f$ we shall be studying has its origins in statistical mechanics. In this context it is only necessary to consider its action on the space of real-valued functions (or interactions) of exponentially decreasing variation $\mathcal{F}_\theta$. Ruelle showed that the spectrum of $\mathcal{L}_f : \mathcal{F}_\theta \to \mathcal{F}_\theta$ ($f \in \mathcal{F}_\theta$) satisfies a Perron-Frobenius type theorem (Theorem 1) [15]. Subsequently other proofs of this, and other related results, have been developed [20], [4], ([17] p. 83). In Section 1, we present a new proof of the existence of a maximal eigenvalue for $\mathcal{L}_f$.

One major application of Ruelle's theorem is the construction of meromorphic extensions for certain generalised zeta functions [16].

It is the purpose of this paper to present a generalisation of this theorem to describe the spectrum of $\mathcal{L}_f$ for complex functions of exponentially decreasing variation (Theorem 2). This subsumes a complex version of the Perron-Frobenius theorem for matrices due to Wielandt (Proposition 1). This new spectral theorem provides a more natural setting for the ingenious techniques developed by Ruelle ([17] pp. 93-95), and enables us to produce extension results for zeta functions (Theorem 3) subsuming those due to Ruelle [16] and Parry and the author ([12], Theorem 1).

In the last two sections we give examples which answer negatively questions raised by Ruelle ([17] p. 173) and Bowen ([1] p. 31) (During the writing of this paper the author discovered that the example in Section 4 was known to Ruelle [18].)

This paper is an offshoot of the joint work of Parry and the author concerning asymptotic estimates for the number of closed orbits for Axiom A flows [12]. Parry has since derived other interesting estimates as a result of applying ideas from analytic number theory to the study of Axiom A flows [11].

I wish to express my gratitude to the S.E.R.C. for their financial support.

I am deeply indebted to Professor William Parry for his help and encouragement throughout the course of this work.

§1. THE RUELLE OPERATOR THEOREM

Let A be an aperiodic zero-one matrix of rank k and define

$$\Sigma_A^+ = \{x \in \prod_0^\infty \{1,2,\ldots,k\} \mid A(x_n,x_{n+1}) = 1\}.$$

The space Σ_A^+ is compact and zero dimensional with respect to the topology with basis consisting of sets of the form $\{x \in \Sigma_A^+ | x_i = z_i, \quad 0 < i < n-1\}$.

The *(one-sided) shift of finite type* $\sigma : \Sigma_A^+ \to \Sigma_A^+$ is the continuous map given by $(\sigma x)_n = x_{n+1}$. Since A is aperiodic σ is (topologically) mixing.

If $f : \Sigma_A^+ \to \mathbb{R}$ is continuous, the *pressure* is defined by

$$P(f) = \sup \{h_\mu(\sigma) + \int f \, d\mu \;\Big|\; \mu \text{ is } \sigma\text{-invariant}\}.$$

This supremum is always attained and the measures for which $P(f) = h_\mu(\sigma) + \int f \, d\mu$ are called *equilibrium states* ([21], p. 224).

Define $\operatorname{var}_n f = \sup\{|f(x) - f(y)| \;\Big|\; x_i = y_i, \; 0 < i < n-1\}$ then for $0 < \theta < 1$ let

$$\|f\|_\theta = \sup\left\{ \frac{\operatorname{var}_n f}{\theta^n} \;\Big|\; n > 0\right\}.$$

In this section our main interest will be in the real Banach space

$$\mathcal{F}_\theta = \{f \in C(\Sigma_A^+) \;\big|\; \|f\|_\theta < \infty\}$$

with norm $|||f|||_\theta = \max\{\|f\|_\infty, \|f\|_\theta\}$.

Given $f \in \mathcal{F}_\theta$ define an operator $\mathcal{L}_f : \mathcal{F}_\theta \to \mathcal{F}_\theta$ by

$$\mathcal{L}_f h(z) = \sum_{\sigma y = z} e^{f(y)} h(y).$$

We now present a proof of the Ruelle operator theorem which does not involve measures. The existence part was suggested by techniques employed by Krasnoselskii in [9] for a different problem. The rest of the proof is a combination of Ruelle ([17] p. 90) and Walters [20].

Theorem 1. (Ruelle).

Let σ be a topologically mixing one-sided shift of finite type and let $f \in \mathcal{F}_\theta$. The operator $\mathcal{L}_f$ has $\exp P(f)$ as a simple eigenvalue (with a positive eigenvector). Furthermore the rest of the spectrum is contained in a disc of radius strictly smaller than $\exp P(f)$.

Proof.

Let Λ denote the $\|\ \|_\infty$-closed set of non-negative continuous functions $g : \Sigma_A^+ \to \mathbb{R}^+$ with $\|g\|_\infty < 1$ and

$$g(x) < g(y) \exp\left(\frac{\theta^n \|f\|_\theta}{1-\theta}\right)$$

whenever $x_i = y_i$, $0 < i < n-1$.

Since

$$|g(x) - g(y)| < \frac{\theta^n}{1-\theta}\|f\|_\theta . \exp\left(\frac{\|f\|_\theta}{1-\theta}\right)$$

Λ is equicontinuous and therefore $\| \quad \|_\infty$-compact by the Arzela-Ascoli theorem.

The continuous map $L_n : \Lambda \to \Lambda$ given by

$L_n g = \mathcal{L}_f(g + 1/n) / \|\mathcal{L}_f(g + 1/n)\|_\infty$ $n > 0$ has a fixed point h_n by the Schauder-Tychonoff theorem. If λ_n denotes $\|\mathcal{L}_f\ (h_n + 1/n)\|_\infty$ then

$$\lambda_n h_n = \mathcal{L}_f(h_n + 1/n) > (\inf h_n + 1/n) e^{-\|f\|_\infty}$$

$$> 0$$

and so $\lambda_n > e^{-\|f\|_\infty}$. If h is a limit point of $\{h_n\}$ then $\mathcal{L}_f h = \lambda h$ where $\lambda = \|\mathcal{L}_f h\|_\infty > 0$.

The eigenfunction h is strictly positive since if $h(x) = 0$ for some x then $\mathcal{L}_f^n h(x) = \sum_{\sigma^n y = x} e^{f^n(y)} h(y) = 0$

(where $f^n(y) = f(y) + f(\sigma y) + \ldots + f(\sigma^{n-1} y)$). This would make h zero on the dense set $\{y | \sigma^n y = x \text{ for some } n > 0\}$ contradicting $\lambda > 0$.

The eigenvalue λ is simple since if $\mathcal{L}_f g = \lambda g$ and $t = \inf\{g/f\}$ then $(tf-g)(x) = 0$ for some x. The preceding argument applied to $tf-g \geq 0$ shows that $tf = g$.

By replacing f by $f + \log h - \log h \circ \sigma - \log \lambda$ we may assume that $\mathcal{L}_f 1 = 1$. The general effect of this change is to scale the spectrum of $\mathcal{L}_f$ by λ.

Since A is aperiodic $A^M > 0$, for some M, given $z, \sigma^n x = x$ choose $\sigma^{n+M} y = z$ with $x_i = y_i$, $0 \leq i \leq n$, then

$$\sum_{\sigma^n x = x} e^{f^n(x)} \leq \sum_{\sigma^{n+M} y = z} \exp\left(f^{n+M}(y) + M\|f\|_\infty + \frac{\|f\|_\theta}{1-\theta}\right) = C\mathcal{L}_f^{n+M} 1(z) = C.$$

Thus $P(f) = \lim \frac{1}{n} \log \sum_{\sigma^n_x = x} e^{f^n(x)} = 0$ by use of the variational principle ([21] p. 218). Thus for our original f, $\log \lambda = P(f)$.

It is easy to show that (for some constant $C > 0$)

$$(1.1) \qquad \|\mathcal{L}_f^n g\|_\theta \leq C \|g\|_\infty + \theta^n \|g\|_\theta$$

for all $g \in \mathcal{F}_\theta$, $n \geq 0$. This means that $\{\mathcal{L}_f^n g\}$ is equicontinuous with respect to $\|\ \|_\infty$ and there exists a limit point ℓ. If we denote $\alpha(g) = \sup\{g(x)\}$ and $\beta(g) = \inf\{g(x)\}$ then

$$\alpha(g) \geq \alpha(\mathcal{L}_f g) \geq \dots \geq \alpha(\ell) = \alpha(\mathcal{L}_f^n \ell) \quad n \geq 0$$

and

$$\beta(g) < \beta(\mathcal{L}_f g) < \dots \quad < \beta(\ell) = \beta(\mathcal{L}_f^n \ell) \quad n > 0.$$

Since σ is mixing the equalities show ℓ is a constant. Furthermore since $\alpha(\ell) = \beta(\ell)$ the sequence $\mathcal{L}_f^n g$ converges uniformly to ℓ.

To remove the maximal eigenvalue consider $\hat{\mathcal{L}}_f$ acting on the quotient space $\mathcal{F}_\theta/\mathbb{R}$. On this space (1.1) becomes

$$\|\mathcal{L}_f^n g\|_\theta < C.\ \mathrm{var}_o g + \theta^n \ \|g\|_\theta\ .$$

Since $\mathrm{var}_o\ \mathcal{L}_f^n g$ converges to zero we have for large n

$$(1.2)\quad \|\mathcal{L}_f^{2n} g\|_\theta < C.\ \mathrm{var}_o\ \mathcal{L}_f^n g + \theta^n\ [C.\mathrm{var}_o g + \theta^n\ \|g\|_\theta].$$
$$< 1.$$

By the uniform compactness of $\{g \mid |||g|||_\theta < 1\}$ we may choose n so that (1.2) holds for all g in this ball. Thus the spectral radius of $\mathcal{L}_f : \mathcal{F}_\theta/\mathbb{R} \to \mathcal{F}_\theta/\mathbb{R}$, denoted by $\rho(\mathcal{L}_f)$, satisfies

$$\rho(\mathcal{L}_f) = \inf\ \{\|\mathcal{L}_f^n\|_\theta^{1/n} \mid n > 0\} < 1.$$

This completes the proof.

§2. A COMPLEX RUELLE OPERATOR THEOREM

A continuous function $f:\Sigma_A^+ \to \mathbb{C}$ is called locally constant if there exists $n > 0$ such that $f(x)$ depends on only the first n places of x, i.e. $f(x) = f(y)$ whenever $x_i = y_i$, $0 < i < n-1$.

The operator $\mathcal{L}_f$ leaves invariant the finite dimensional subspace of $C(\Sigma_A^+, \mathbb{C})$ with base vectors

$$\delta_{x_0,\ldots,x_{n-2}}(z) = \begin{cases} 1 & x_i = z_i,\ 0 < i < n-2 \\ 0 & \text{otherwise.} \end{cases}$$

It is always possible to reduce locally constant functions to the case $n = 2$ by considering $x \in \Sigma_A^+$ as the sequence $(x_m \ldots x_{m+n-2})_{m=0}^{\infty}$ in a shift space whose symbols are words of length $n-1$.

For $n = 2$ $\mathcal{L}_f$ can be represented by a matrix $M = [A(x_0,x_1).\exp f(x_0,x_1)]$. When f is real-valued the eigenvalues of M are described by the Perron-Frobenius theorem. More generally define M_+ to be the positive matrix with entries $|M(x_0,x_1)|$. The following result is due to Wielandt ([6] p. 57).

Proposition 1.

The eigenvalues of M have moduli less than or equal to the maximal eigenvalue β for M_+. If βe^{ia} is an eigenvalue of M (for some $0 < a < 2\pi$) then M takes the form $M = e^{ia} DM_+D^{-1}$ where D is a diagonal matrix with diagonal entries of unit modulus.

Ruelle's theorem (Theorem 1) can be viewed as a generalisation of the Perron-Frobenius theorem. In this section we present an analogous extension of Wielandt's result.

The space $\mathcal{F}_\theta^{\mathbb{C}} = \{f \in C(\Sigma_A^+,\mathbb{C}) \mid \|f\|_\theta < \infty\}$ is a complex Banach space with norm $|||f|||_\theta = \max\ \{\|f\|_\theta, \|f\|_\infty\}$ (Here $\|\ \ \|_\theta$ has the same definition as in Section 1). If $f = u + iv \in \mathcal{F}_\theta^{\mathbb{C}}$ then $u,v \in \mathcal{F}_\theta$ and we freely assume that $\mathcal{L}_u 1 = 1$, as in the proof of Theorem 1.

For $g \in C(\Sigma_A^+)$ let Γ_g be the multiplicative group generated by $\langle \exp g^n(x) \mid \sigma^n x = x\rangle$ (where $g^n(x) = g(x) + g(\sigma x) + \dots + g(\sigma^{n-1}x))$ [13], [12].

<u>Proposition 2</u>.

For $f = u + iv \in \mathcal{F}_\theta^{\mathbb{C}}$ and $0 \le a < 2\pi$ the following are equivalent:

(i) Γ_{v-a} is generated by a power of $e^{2\pi}$

(ii) $\lambda_a \equiv \exp\ [ia + P(u)]$ is an eigenvalue for $\mathcal{L}_f$

(iii) There exists $\omega \in C(\Sigma_A^+)$ such that

$$v-a + \omega \circ \sigma - \omega \in C(\Sigma_A^+,\ 2\pi\mathbb{Z}).$$

<u>Proof</u> (i) $\Rightarrow$ (ii) Choose $x \in \Sigma_A^+$ with a dense orbit and define h on $\{\sigma^n x \mid n \ge 0\}$ by $h(\sigma^n x) = \exp i\ (v-a)^n(x)$. This extends to

an element of $\mathcal{F}_\theta^{\mathbb{C}}$ [10]. Furthermore h satisfies $h(\sigma z) = h(z) \exp i(v-a)(z)$ and consequently $\mathcal{L}_f h = e^{ia}h$.

(ii) $\Rightarrow$ (iii) Since $\mathcal{L}_f h = e^{ia}h$ may be expressed as

$$(2.1) \qquad \sum_{\sigma y = x} e^{u(y)} . e^{i(v-a)(y)} h(y) = h(x)$$

if $|h(x)|$ is maximal then so is $|h(y)|$ when $\sigma y = x$. Because σ is mixing h is of constant modulus. Thus (2.1) represents a convex combination of points on a circle which also lies on the circle. From this we deduce

$$e^{i(v-a)(y)} h(y) = h(\sigma y)$$

or equivalently $v-a + \arg h - \arg h \circ \sigma \in C(\Sigma_A^+, 2\pi\mathbb{Z})$.

(iii) $\Rightarrow$ (i) This is immediate.

If $f \in \mathcal{F}_\theta^{\mathbb{C}}$ satisfies one, and hence all, of the above conditions we call it an *a-function*. For example, the functions in $\mathcal{F}_\theta$ are all a-functions with a = 0.

If f is not an a-function (for any a) then we call it *regular*.

If $f = u + iv$ where $v-a + \omega \circ \sigma - \omega \in C(\Sigma_A^+, 2\pi\mathbb{Z})$ then $\mathcal{L}_f = e^{ia} \Delta(e^{i\omega}) \mathcal{L}_u \Delta(e^{-i\omega})$ where $\Delta(h)$ denotes the operator that multiplies functions by h. Therefore the spectrum of $\mathcal{L}_f$ is precisely the spectrum of $\mathcal{L}_u$ rotated through an angle a. By Theorem 1:

an element of $\mathcal{F}_\theta^{\mathbb{C}}$ [10]. Furthermore h satisfies $h(\sigma z) = h(z) \exp i(v-a)(z)$ and consequently $\mathcal{L}_f h = e^{ia} h$.

(ii) $\Rightarrow$ (iii) Since $\mathcal{L}_f h = e^{ia} h$ may be expressed as

$$(2.1) \qquad \sum_{\sigma y = x} e^{u(y)} . e^{i(v-a)(y)} h(y) = h(x)$$

if $|h(x)|$ is maximal then so is $|h(y)|$ when $\sigma y = x$. Because σ is mixing h is of constant modulus. Thus (2.1) represents a convex combination of points on a circle which also lies on the circle. From this we deduce

$$e^{i(v-a)(y)} h(y) = h(\sigma y)$$

or equivalently $v-a + \arg h - \arg h \circ \sigma \in C(\Sigma_A^+, 2\pi\mathbb{Z})$.

(iii) $\Rightarrow$ (i) This is immediate.

If $f \in \mathcal{F}_\theta^{\mathbb{C}}$ satisfies one, and hence all, of the above conditions we call it an *a-function*. For example, the functions in $\mathcal{F}_\theta$ are all a-functions with $a = 0$.

If f is not an a-function (for any a) then we call it *regular*.

If $f = u + iv$ where $v-a + \omega \circ \sigma - \omega \in C(\Sigma_A^+, 2\pi\mathbb{Z})$ then $\mathcal{L}_f = e^{ia} \Delta(e^{i\omega}) \mathcal{L}_u \Delta(e^{-i\omega})$ where $\Delta(h)$ denotes the operator that multiplies functions by h. Therefore the spectrum of $\mathcal{L}_f$ is precisely the spectrum of $\mathcal{L}_u$ rotated through an angle a. By Theorem 1:

<u>Proposition 3</u>.

If $f = u + iv$ is an a-function then λ_a is a simple eigenvalue for $\mathcal{L}_f$ and the rest of the spectrum is contained in a disc of radius strictly smaller than $|\lambda_a| = \exp P(u)$.

An immediate corollary is that $f \in \mathcal{F}_\theta^{\mathbb{C}}$ can be an a-function for at most one a $(0 \leq a < 2\pi)$. ♦.

For any $f = u + iv$ (with $\mathcal{L}_u 1 = 1$) the following extension of (1.1) is true

$$(2.2) \qquad \|\mathcal{L}_f^n h\|_\theta \leq C.\ \|h\|_\infty + \theta^n \ \|h\|_\theta \quad n \geq 0,\ h \in \mathcal{F}_\theta^{\mathbb{C}}.$$

The operator norm satisfies $|||\mathcal{L}_f^n|||_\theta \leq C + 1$ and we have an upper bound on the spectral radius

$$\rho(\mathcal{L}_f) = \inf \ \{ |||\mathcal{L}_f^n|||_\theta^{\ 1/n} \ | n \geq 1\} \leq 1.$$

We now show that when f is regular the spectrum of $\mathcal{L}_f$ (denoted $sp(\mathcal{L}_f)$) is disjoint from the unit circle.

Choose a point e^{it} on the circle then for $|||h|||_\theta \leq 1$ write $h_N = \frac{1}{N} \sum_{n=0}^{N-1} \mathcal{L}_{f-it}^n h \quad (N \geq 1)$. By (2.2) h_N is contained in the uniformly compact set $\{g \ \big| \ |||g|||_\theta \leq C + 1\}$. When f is regular $\|h_N\|_\infty$ must tend to zero since any non-zero limit point of $\{h_N\}$ would be an eigenvector for $\mathcal{L}_f$ with eigenvalue e^{it}.

<u>Proposition 3</u>.

If $f = u + iv$ is an a-function then λ_a is a simple eigenvalue for $\mathcal{L}_f$ and the rest of the spectrum is contained in a disc of radius strictly smaller than $|\lambda_a| = \exp P(u)$.

An immediate corollary is that $f \in \mathcal{F}_\theta^{\mathbb{C}}$ can be an a-function for at most one a $(0 < a < 2\pi)$.

For any $f = u + iv$ (with $\mathcal{L}_u 1 = 1$) the following extension of (1.1) is true

$$(2.2) \qquad \|\mathcal{L}_f^n h\|_\theta < C. \|h\|_\infty + \theta^n \ \|h\|_\theta \quad n > 0,\ h \in \mathcal{F}_\theta^{\mathbb{C}}.$$

The operator norm satisfies $|||\mathcal{L}_f^n|||_\theta < C + 1$ and we have an upper bound on the spectral radius

$$\rho(\mathcal{L}_f) = \inf \ \{ |||\mathcal{L}_f^n|||_\theta^{\ 1/n} \ | n > 1\} < 1.$$

We now show that when f is regular the spectrum of $\mathcal{L}_f$ (denoted $sp(\mathcal{L}_f)$) is disjoint from the unit circle.

Choose a point e^{it} on the circle then for $|||h|||_\theta < 1$ write $h_N = \frac{1}{N} \sum_{n=0}^{N-1} \mathcal{L}_{f-it}^n h \quad (N > 1)$. By (2.2) h_N is contained in the uniformly compact set $\{g \ \big| \ |||g|||_\theta < C + 1\}$. When f is regular $\|h_N\|_\infty$ must tend to zero since any non-zero limit point of $\{h_N\}$ would be an eigenvector for $\mathcal{L}_f$ with eigenvalue e^{it}.

For $k > 0$,

$$||h_N||_\theta < ||h_N - \mathcal{L}^k_{f-it} h_N||_\theta + ||\mathcal{L}^k_{f-it} h_N||_\theta$$

where we have estimates

$$||h_N - \mathcal{L}^k_{f-it} h_N||_\theta = ||\frac{1}{N} \sum_{n=0}^{k-1} \mathcal{L}^n_{f-it} h - \mathcal{L}^{n+N}_{f-it} h||_\theta$$

$$< \frac{2k}{N} (C + 1)$$

and

$$||\mathcal{L}^k_{f-it} h_N||_\theta < C \, ||h_N||_\infty + \theta^k \, ||h_N||_\theta$$

$$< C.||h_N||_\infty + \theta^k (C + 1).$$

If we take $k = [N^{\frac{1}{2}}]$ then these bounds show that $\overline{\lim} \; ||h_N||_\theta = 0$. By the uniform compactness of $\{h \;\big|\; |||h|||_\theta < 1\}$ we can choose N such that $|||h_N|||_\theta < 1$ for all h in this ball. Since the spectral radius of an operator is smaller than its norm

$$\rho\left(\frac{1}{N} \sum_{n=0}^{N-1} \mathcal{L}^n_{f-it}\right) < |||\frac{1}{N} \sum_{n=0}^{N-1} \mathcal{L}^n_{f-it}|||_\theta < 1.$$

Thus 1 is not an element of $sp\left(\frac{1}{N} \sum_{n=0}^{N-1} \mathcal{L}^n_{f-it}\right)$. However we know from the spectral mapping theorem ([19] p. 263) that

$$sp\left(\frac{1}{N} \sum_{n=0}^{N-1} \mathcal{L}^n_{f-it}\right) = \left\{\frac{1}{N} \sum_{n=0}^{N-1} \lambda^n \mid \lambda \in sp(\mathcal{L}_{f-it})\right\}.$$

Thus 1 cannot be an element of $sp(\mathcal{L}_{f-it}) = e^{-it}\, sp(\mathcal{L}_f)$ or equivalently e^{it} is not in the spectrum of $\mathcal{L}_f$. Since e^{it} was chosen arbitrarily and $sp(\mathcal{L}_f)$ is closed we have the following.

Theorem 2.

Let σ be a topologically mixing one-sided shift of finite type and let $f = u + iv \in \mathcal{F}_\theta^{\mathbb{C}}$

(i) If f is an a-function then $\lambda_a = \exp\,[ia + P(u)]$ is a simple eigenvalue for $\mathcal{L}_f$ and the rest of the spectrum is contained in a disc of radius strictly smaller than $|\lambda_a| = \exp P(u)$

(ii) If f is regular then the spectrum of $\mathcal{L}_f$ is contained in a disc of radius strictly smaller than $\exp P(u)$.

It is possible to formulate a proof of part (ii) closer to the proof of Theorem 1 by proceding along the lines of Propositions 13 and 14 in [12].

An important feature of the above theorem is that the type of spectrum for $\mathcal{L}_f$ is determined by $\mathfrak{I}(f)$ and the size of the spectrum is given by $\mathfrak{R}(f)$ (and also $\mathfrak{I}(f)$)

We can obtain a *lower bound* for the spectral radius $\rho(\mathcal{L}_f)$ in the regular case. If $f = u + iv$ then $\mathcal{L}_u^n = \mathcal{L}_f^n \Delta(e^{-iv^n})$ and

$$\exp P(u) = \rho(\mathcal{L}_u) < \rho(\mathcal{L}_f).\overline{\lim}\ \||\Delta(e^{-iv^n})\||_\theta^{1/n} .$$

It is simple to show $\overline{\lim}\ \||\Delta(e^{-iv^n})\||_\theta^{1/n} < 1/\theta$ and so $\rho(\mathcal{L}_f) > \theta.\exp P(u)$.

Proposition 2(i) shows that a necessary condition for $u + iv$ to be an a-function (for some a) is that Γ_v^* should be of rank at most two. But functions satisfying this condition can easily be approximated by functions which do not. This makes the family of regular functions dense in $\mathcal{F}_\theta^{\mathbb{C}}$. Furthermore, by Theorem 2 and upper semicontinuity of $f \mapsto \rho(\mathcal{L}_f)$ this family is also open.

§3. EXTENDING THE ZETA FUNCTION

Given $f \in \mathcal{F}_\theta^{\mathbb{C}}$ define a zeta function by

$$\zeta(f) = \exp \sum_{n=1}^{\infty} \frac{1}{m} \sum_{\sigma^m x = x} \exp f^m(x)$$

where $f^m(x) = f(x) + f(\sigma x) + \dots + f(\sigma^{m-1} x)$.

If $f = u + iv$ with $P(u) < 0$ then

$$\overline{\lim_m} \Big| \sum_{\sigma^m x = x} \exp f^m(x) \Big|^{1/m} < \overline{\lim_m} \Big| \sum_{\sigma^m x = x} \exp u^m(x) \Big|^{1/m}$$

$$= \exp P(u) < 1.$$

Since convergence is uniform in a neighbourhood of f it follows that ζ is non-zero and analytic on $\{f \,|\, \mathbb{R}(f) < 0\}$ ([17] p. 100).

We now consider the cases where $P(u) = 0$.

Proposition 4. If $f = u + iv \in \mathcal{F}_\theta^{\mathbb{C}}$ is regular with $P(u) = 0$ then ζ is analytic and non-zero in a neighbourhood of f.

Proof.

Choose $\theta' > \theta$ then $f \in \mathcal{F}_\theta^{\mathbb{C}} \subseteq \mathcal{F}_{\theta'}^{\mathbb{C}}$. For $n > 0$ define locally constant functions

$$u_n(x) = \sup \{u(z) \,|\, x_i = z_i, \quad 0 < i < n-1\}$$
$$v_n(x) = \sup \{v(z) \,|\, x_i = z_i, \quad 0 < i < n-1\}$$

and let $f_n(x) = u_n(x) + iv_n(x)$.

This enables us to write

$$(3.1)\qquad \zeta(f) = \exp \sum_{m=1}^{\infty} \frac{1}{m} \sum_{\sigma^m x = x} \exp f^m(x) - \exp f_n^m(x)$$

$$\times \exp \sum_{m=1}^{\infty} \frac{1}{m} \sum_{\sigma^m x = x} \exp f_n^m(x).$$

By applying Theorem 2 to $\mathcal{L}_f : \mathcal{F}_{\theta'}^{\mathbb{C}} \to \mathcal{F}_{\theta'}^{\mathbb{C}}$ there exists $0 < \beta < 1$ such that spectral radius $\rho(\mathcal{L}_f) < \beta$. Since $g \mapsto \rho(\mathcal{L}_g)$ is upper semicontinuous on $\mathcal{F}_{\theta'}^{\mathbb{C}}$ and

$$|||f - f_n|||_{\theta'} < ||f||_{\theta}\left(\frac{\theta}{\theta'}\right)^n$$

it follows that $\rho(\mathcal{L}_{f_n}) < \beta$ for large enough n. If ε is chosen sufficiently small then $\rho(\mathcal{L}_{g_n}) < \beta$ holds uniformly on $D = \{g \mid |||g - f|||_{\theta} < \varepsilon\}$.

Let $\lambda_1,\ldots,\lambda_{N(n)}$ be the eigenvalues of $\mathcal{L}_{f_n}$ acting on the finite dimensional invariant subspace of Section 2 then $|\lambda_j| < \beta$, $1 \le j \le N(n)$.

Furthermore

$$\sum_{\sigma^m x = x} \exp f_n^m(x) = \text{trace } \mathcal{L}_{f_n}^m = \lambda_1^m + \ldots + \lambda_{N(n)}^m$$

where $N(n) \le k^n$ (k is the dimension of A). Choose α satisfying $\beta k^{\alpha} < 1$ and take $n = [m\alpha]$. Then

$$\Big| \sum_{\sigma^m x = x} \exp f_n^m(x) \Big|^{1/m} < (\beta^m k^n)^{1/m} < \beta k^\alpha < 1$$

for sufficiently large m. Since this holds uniformly on D

$$\sum_{m=1}^{\infty} \frac{1}{m} \sum_{\sigma^m x = x} \exp g_n^m(x)$$

is analytic for $g \in D$.

From the definitions of u_n and v_n

$$(3.2) \qquad \|u_n^m - u^m\|_\infty < m\,\|u_n - u\|_\infty < m\theta^n\, \|u\|_\theta$$

$$\|v_n^m - v^m\|_\infty < m\theta^n\, \|v\|_\theta\ .$$

Since

$$\sum_{\sigma^m x = x} e^{f^m(x)} - e^{f_n^m(x)}$$

$$= \sum_{\sigma^m x = x} \left[e^{u^m(x)} - e^{u_n^m(x)}\right] e^{iv^m(x)} + \sum_{\sigma^m x = x} e^{u_n^m(x)} \left[e^{iv^m(x)} - e^{iv_n^m(x)}\right]$$

it follows from (3.2) that

$$\overline{\lim}\ \Big| \sum_{\sigma^m x = x} \exp f^m(x) - \exp f_n^m(x) \Big|^{1/m} < \theta^\alpha \exp P(u).$$

If we choose $\theta^\alpha < \Phi < 1$ and ε sufficiently small then for large m

$$\Big| \sum_{\sigma^m x = x} \exp g^m(x) - \exp g_n^m(x) \Big|^{1/m} < \Phi$$

uniformly on D and so

$$\sum_{m=1}^{\infty} \frac{1}{m} \sum_{\sigma^m x = x} \exp g^m(x) - \exp g_n^m(x)$$

is analytic on this disc.

This completes the proof.

When f is an a-function then $\rho(\mathcal{L}_f) = \exp P(\mathfrak{R}(f))$. However the isolated eigenvalue λ_a can be dealt with using perturbation theory. In a neighbourhood of f the operator $\mathcal{L}_g$ still has an isolated eigenvalue β ([3] p. 587). This leads to a natural definition of the complex pressure (in a neighbourhood of an a-function) as $P(g) = \log \beta$.

By developing an approach due to Ruelle ([17] pp. 93-95), Parry has proved the following result ([11], Proposition 3).

<u>Proposition 5</u>.

If $f \in \mathcal{F}_\theta$ and $P(f) = 0$ then there exists $\varepsilon > 0$ such that P extends to an analytic function in

$$D = \{g \;\big|\; |||f-g|||_\theta < \varepsilon\} \text{ and}$$

$$\sum_{m=1}^{\infty} \frac{e^{iam}}{m} \Big(\sum_{\sigma^m x = x} g^m(x) - e^{mP(g)} \Big)$$

converges uniformly in D(for any $a \in \mathbb{R}$)

Propositions 4 and 5 together give the following result (The version for two-sided shifts is Theorem 1 in [11]).

<u>Theorem 3</u>.

Let $f = u + iv \in \mathcal{F}_\theta^{\mathbb{C}}$.

(i) If $P(u) < 0$ or f is regular with $P(u) = 0$ then ζ is non-zero and analytic in a neighbourhood of f.

(ii) If f is an a-function with $P(u) = 0$ then ζ has a non-zero analytic extension to a set $\{g \mid |||g-f|||_\theta < \varepsilon, P(g) \neq 0\}$ given by

$$\zeta(g) = \frac{1}{1-e^{P(g)}} \exp \sum_{m=1}^{\infty} \frac{1}{m} \Big(\sum_{\sigma^m x = x} g^m(x) - e^{mP(g)} \Big).$$

The above theorem extends a result of Ruelle [16], ([17] pp. 100-101).

§4. A COUNTER-EXAMPLE TO RUELLE'S QUESTION

Let $\Sigma_A = \{x \in \prod_{-\infty}^{\infty} \{1,\dots,k\} | A(x_n,x_{n+1}) = 1\}$ then $\sigma:\Sigma_A \to \Sigma_A$ given by $(\sigma x)_n = x_{n+1}$ is a *(two-sided) shift of finite type*. Let $f:\Sigma_A \to \mathbb{R}^+$ be a strictly positive continuous function for which there exists $0 < \theta < 1$, $C > 0$ satisfying $|f(x) - f(y)| \leq C\theta^n$ whenever $x_i = y_i$, $|i| \leq n-1$. Define

$$\Sigma_A^f = \{(x,t)|0 \leq t \leq f(x)\}$$

where $(x,f(x))$ and $(\sigma x,0)$ are identified.

The f *suspension* $\sigma_t^f : \Sigma_A^f \to \Sigma_A^f$ is the flow defined by $\sigma_t^f(x,s) = (x,t+s)$ with appropriate identifications. Thus σ^f can be interpreted as flowing vertically under the graph of f.

The flow σ^f is *(topologically) weak mixing* if the rank of $\Gamma_f = \langle \exp f^n(x)|\sigma^n x = x\rangle$ is greater than one [12].

The *topological entropy* of σ^f is the unique $h \in \mathbb{R}^+$ satisfying $P(-hf) = 0$ [12].

The zeta function associated with σ^f is

$$Z(s) = \zeta(-sf) = \exp \sum_{m=1}^{\infty} \frac{1}{m} \sum_{\sigma^m x = x} \exp - sf^m(x).$$

In [12] Parry and the author partially answered a question of Ruelle ([17] p. 173) by showing that $Z(s)$ has an analytic extension to a neighbourhood of $\{s|\mathfrak{R}(s) \geq h\}$, except for a simple pole at $s = h$. We shall now complete this analysis by presenting a flow for which $Z(s)$ is not analytic on any strip $h-\delta < \mathfrak{R}(s) < h$.

Let $\sigma:\Sigma_A \to \Sigma_A$ be a full shift on two symbols $\{1,2\}$. Choose $0 < p < \frac{1}{2}$, $p + q = 1$, and define a locally constant function f by

$$f(x) = -\log p \quad \text{if } x_0 = 1$$

$$\qquad\quad -\log q \quad \text{if } x_0 = 2.$$

If $P^s = \begin{pmatrix} p^s & q^s \\ p^s & q^s \end{pmatrix}$ where $s \in \mathbb{C}$ then

$$\sum_{\sigma^m x = x} \exp -s f^m(x) = \operatorname{trace} (P^s)^m = (p^s + q^s)^m.$$

Thus

$$Z(s) = \exp \sum_{m=1}^{\infty} \frac{1}{m} (p^s + q^s)^m$$

$$= \exp - \log (1 - p^s - q^s) = \frac{1}{1-p^s-q^s}$$

and the poles for $Z(s)$ are the solutions to $p^s + q^s = 1$. In particular the first part of the question shows $h = 1$ and for σ^f to be weak mixing we require $\log p/\log q$ to be irrational.

Let $\varepsilon > 0$ satisfy $p^{-\varepsilon} - q^{-\varepsilon} = 1$ then the poles are contained in the strip $-\varepsilon < \mathfrak{R}(s) < 1$. If $-\varepsilon < \sigma < 1$ then zero is a limit point of $\{p^s + q^s - 1 | \mathfrak{R}(s) = \sigma\}$. Since

$p^s + q^s - 1$ is an analytic almost periodic function it has a zero in every vertical strip containing σ ([2], p. 75). We conclude that the poles $\{\sigma_n + it_n\}$ for Z are distributed with $\{\sigma_n\}$ dense in the interval $[-\epsilon,1]$. (In fact sharper estimates about the distribution of poles are possible (cf. [8])).

§5. A COUNTER-EXAMPLE TO BOWEN'S QUESTION

In [5] Gallovotti gave an example of a suspension for which the corresponding zeta function has an essential singularity at $s_1 < 0$. Bowen asked whether the zeta function for flows could always be extended to $s = 0$ ([1], p. 31). In this section we give an example where this is not the case. In fact it is possible to construct a suspension with an essential singularity at $s_0 > 0$.

Let $\sigma_n : \Sigma_n \to \Sigma_n$ be a full shift on n-symbols and let $\{\beta_k\}$ be a convergent sequence with limit β. For $n = 3$ define $g \in C(\Sigma_3)$ by

$$
\begin{aligned}
g(z) &= \beta_k \text{ if } z_k = 2,\ z_i \in \{1,3\} \quad 0 < i < k-1 \\
&= \beta \quad \text{if } z_i \in \{1,3\} \quad i > 0.
\end{aligned}
$$

Let $\sigma^m z = z$ and assume the cycle $(z_0,\ldots,z_{m-1})$ contains disjoint blocks of 1s and 3s of lengths $k_1,\ldots,k_r$ with $k_1 + \ldots + k_r = N$. Then

$$g^m(z) = (m-N)\beta_0 + \sum_{p=1}^{r} (\beta_1 + \ldots + \beta_{k_p}).$$

Thus $g^m(z)$ is independent of the 2^N possible combinations of 1s and 3s.

For $n = 2$ define $f \in C(\Sigma_2)$ by

$$
\begin{aligned}
f(x) &= \beta_k + \log 2 \quad \text{if } x_k = 2, \quad x_i = 1 \\
&\qquad\qquad\qquad\quad 0 < i < k-1, \quad (k \neq 0) \\
&= \beta_0 \qquad\qquad\; \text{if } x_0 = 2 \\
&= \beta + \log 2 \quad\; \text{if } x_i = 1,\ i > 0.
\end{aligned}
$$

The functions f and g are related by

$$\sum_{\sigma^m x = x} \exp f^m(x) = \sum_{\sigma^m z = z} \exp g^m(z)$$

and so $\zeta(f) = \zeta(g)$.

The function f is similar to the Fisher potential used by Gallovotti [5].

Define a locally constant function f_N by replacing β_k by β for $k > N$. The zeta function $\zeta(f_N)$ can be simply calculated. Define

$$
P_N = \begin{bmatrix}
e^{\beta_0} & 2e^{\beta_1} & \cdots & 2e^{\beta_{N-1}} & 2e^{\beta} \\
e^{\beta_0} & & & \bigcirc & 0 \\
 & 2e^{\beta_1} & & & \vdots \\
\bigcirc & & \ddots & & 0 \\
 & & & 2e^{\beta_{N-1}} & 2e^{\beta}
\end{bmatrix}
$$

then by ([14], p. 82)

$$(5.1) \qquad 1/\zeta\,(f_N) = \det\,(I - P_N)$$

$$= (1 - 2e^{\beta})(1 - \sum_{n=0}^{N-1} 2^n e^{\beta_0+\ldots+\beta_n})$$

$$- 2^N \quad e^{\beta+\beta_0 + \ldots + \beta_{N-1}}.$$

Assume that $g \in \mathcal{F}_\theta$ and $g > 0$. By replacing β_k by $-s\beta_k$ (and β by $-s\beta$) we have from (5.1) and Section 3

$$(5.2) \quad 1/Z(s) = 1/\zeta(-sf) = (1-2e^{-s\beta})(1 - \sum_{n=0}^{\infty} 2^n e^{-s(\beta_0+\ldots+\beta_n)})$$

(for $\Re(s)$ large).

Following Gallovotti we set

$$\beta_m = -\log\left(\frac{1+\theta^m/m}{1+\theta^{m-1}/m-1}\right) + C \qquad m \geq 2$$

$$= -\log\,(1 + \theta) + C \qquad m = 1$$

$$= C \qquad m = 0$$

and $\alpha = C$ (where $C > 0$ is chosen to make $\beta_k > 0$.)

From (5.2)

$$1/Z(s) = (1-2e^{-sC})(1 - \sum_{m=1}^{\infty} (1+\theta^m/m)^s (2e^{-sC})^{m+1}/2 - e^{-sC})$$

Thus the entropy of σ^g is the solution $h > 0$ to

$$1 = \sum_{m=1}^{\infty} (1 + \theta^m/m)^h (2e^{-hC})^{m+1}/2 + e^{-hC}$$

Z(s) has a meromorphic extension to s = h given by
$1/Z(s) = (1-2e^{-sC})(1-e^{-sC}[F(s)+1]) - 2e^{-2sC}$ where

$$F(s) = \sum_{m=1}^{\infty} (2e^{-sC})^m [(1+\theta^m/m)^s - 1].$$

For $0 < s < h$ there exists $B, D > 0$ such that

$$B.s.\frac{\theta^m}{m} < (1+\theta^m/m)^s - 1 < D.s.\frac{\theta^m}{m}.$$

Thus

$$B.\log(1 - 2e^{-sC}\theta) < \frac{F(s)}{s} < D.\log(1-2e^{-sC}\theta).$$

Consider $s_0 = 1/C \log 2\theta$. If $s_0 > 0$ (or equivalently $\theta > \frac{1}{2}$) then as s approaches s_0 from above $|F(s)|$ is unbounded but $(s-s_0)F(s)$ tends to zero. If $s_0 = 0$ (or equivalently $\theta = \frac{1}{2}$) then as s approaches zero from above $|sF(s)|$ is unbounded but $s^2F(s)$ tends to zero. We conclude that in either case s_0 is an essential singularity.

<u>Remark</u>

Hofbauer used the Fisher potential to produce examples of functions with two equilibrium states (one a single atom) [7]. The type of functions studied in this section give examples with two *non-atomic* equilibrium states (one with support a Cantor set).

<u>Remark</u>

Our example extends in a natural way to suspensions over Σ_n for $n > 2$. This enables us to give an example with an essential singularity s_0 arbitrarily close to $h(\sigma) = 1$.

REFERENCES

[1] R. Bowen, *On Axiom A diffeomorphisms*, Amer. Math. Soc. Regional Conf. Proc., No. 35, 1978.

[2] C. Corduneanu, *Almost periodic functions*, Interscience, New York, 1968.

[3] N. Dunford and J.T. Schwartz, *Linear Operators*, Part I, Interscience, New York, 1958.

[4] P. Ferrero and B. Schmitt, Ruelle's Perron-Frobenius theorem and projective metrics, Colloquia Mathematica Societatis János Bolyai, *27* (1979) 333-336.

[5] G. Gallovotti, Funzioni zeta ed insiemi basilar, Accad. Lincei. Rend. Sc. fismat. e nat., *61* (1976) 309-317.

[6] F.R. Gantmacher, *The theory of matrices*, vol. II, Chelsea, New York, 1974.

[7] F. Hofbauer, Examples of the nonuniqueness of the equilibrium state, Trans. Amer. Math. Soc., *228* (1977) 223-241.

[8] B. Jessen and H. Tornhave, Mean motions and almost periodic functions, Acta Math., *77* (1945) 137-279.

[9] M. Krasnoselskii, *Positive solutions of operator equations*, P. Noordhoff, Groningen, 1964.

[10] A.N. Livsic, Cohomology of dynamical systems, Math. USSR Izvestiza, *6* (1972) 1276-1301.

[11] W. Parry, Bowen's equidistribution theory and the Dirichlet density theorem, (to appear)

[12] W. Parry and M. Pollicott, An analogue of the prime number theorem for closed orbits of Axiom A flows (to appear).

[13] W. Parry and K. Schmidt, Natural coefficients and invariants for Markov shifts (to appear).

[14] W. Parry and S. Tuncel, *Classification problems in ergodic theory*, L.M.S. Lecture Notes *67*, C.U.P., Cambridge, 1982.

[15] D. Ruelle, Statistical mechanics of a one-dimensional lattice gas, Commun. Math. Phys., *9* (1968) 267-278.

[16] D. Ruelle, Generalised zeta functions for Axiom A basic sets, Bull. Amer. Math. Soc., *82* (1976) 153-156.

[17] D. Ruelle, *Thermodynamic Formalism*, Addison-Wesley, Reading, 1978.

[18] D. Ruelle, Flows which do not exponentially mix. Comptes rendus, *296* Série I, No. 4 (1983) 191-194.

[19] A.E. Taylor, An introduction to functional analysis, Wiley, New York, 1964.

[20] P. Walters, Ruelle's operator theorem and g-measures, Trans. Amer. Math. Soc., *214* (1975) 375-387.

[21] P. Walters, *An introduction to ergodic theory*, G.T.M. *79*, Springer, 1981.

An analogue of the prime number theorem for closed orbits of Axiom A flows

By William Parry and Mark Pollicott

ANNALS OF MATHEMATICS

In Section 1 we recall some known results

suspension flows. As an application we relate

Annals of Mathematics, 118 (1983), 573–591

An analogue of the prime number theorem for closed orbits of Axiom A flows

By William Parry and Mark Pollicott

Abstract

For an Axiom A flow restricted to a basic set we extend the zeta function to an open set containing $\mathscr{R}(s) \geq h$ where h is the topological entropy. This enables us to give an asymptotic formula for the number of closed orbits by adapting the Wiener-Ikehara proof of the prime number theorem.

For a geodesic flow on a d dimensional compact manifold M^d of constant negative curvature ($\kappa = -1$), Margulis [9] announced a proof that the number of closed orbits τ with least period $\lambda(\tau)$ not exceeding x is asymptotic to $e^{(d-1)x}/(d-1)x$. (See also Hejhal [7] who provides a proof based on the Selberg zeta function.) This is a special case of the theorem which, according to Alexeev and Jacobson [2], appears in Margulis's dissertation and which shows that

$$\#\{\tau\colon \lambda(\tau) \leq x\} \sim e^{hx}/hx$$

for Anosov flows, where h is the topological entropy. (Presumably the flow is assumed to be mixing.)

For Axiom A flows Bowen, [3] (see also Sinai [14]), proved the existence of positive constants A, B such that

$$A \cdot e^{hx}/hx \leq \#\{\tau\colon \lambda(\tau) \leq x\} \leq B \cdot e^{hx}/hx,$$

and conjectured Margulis's more precise result for (topologically weak-mixing) flows of this type. Equivalently he conjectured that

$$\pi(x) = \#\{\tau\colon N_h(\tau) \leq x\} \sim x/\log x \tag{0.1}$$

where $N_h(\tau) = e^{\lambda(\tau)h}$.

Our aim is to prove (0.1) for (basic sets of) topologically weak-mixing Axiom A flows (Theorem 2). A modified asymptotic formula for Axiom A flows which are not topologically weak-mixing is also established.

Sarnak and Woo (cf. [13]) obtain very precise asymptotic estimates for $\pi(x)$ for geodesic flows on non-compact manifolds with finite volume, a case not covered by our results in their present form.

As one would expect in this area the proof of (0.1) depends very heavily on the work of Bowen [5]. We have stated the result in the form (0.1) to emphasise the analogy with the prime number theorem, and in fact our proof follows the main lines of the Wiener-Ikehara proof of that theorem once the relevant properties of an appropriate zeta function have been gathered. Especially relevant is the analytic extension of the zeta function to an open domain containing $\mathscr{R}(s) \geq h$, $s \neq h$ (Theorem 1). (When the flow is not topologically weak-mixing, points $h + nit_0$, $n \in \mathbb{Z}$, have to be excluded, for some $t_0 > 0$.) The possibility of this extension almost answers a question raised by Ruelle in [12]. It will be clear that we utilise many of Ruelle's thermodynamic ideas to establish properties of the zeta function.

We wish to thank C. Series, R. Spatzier and P. Walters for various enlightening remarks and references in connection with our work. Our special thanks go to D. Ruelle who helped us with the proof of his theorem (Proposition 11).

This work is a generalisation to Axiom A flows of results achieved for suspensions, by locally constant functions, of shifts of finite type (cf. [11]).

1. Axiom A flows

Let M be a compact Riemannian manifold and let ϕ_t: $M \to M$ ($t \in \mathbb{R}$) be a C^1 flow. A closed invariant set $\Omega \subset M$ without fixed points is *hyperbolic* if the tangent bundle restricted to Ω is a Whitney sum

$$T_\Omega M = E + E^s + E^u$$

of three $T\phi_t$ invariant sub-bundles, where E is the one dimensional bundle tangent to the flow and E^s, E^u are exponentially contracting and expanding, respectively:

(i) $$\|T\phi_t(v)\| \leq Ke^{-\lambda t}\|v\| \quad \text{for } v \in E^s,\ t \geq 0,$$

(ii) $$\|T\phi_{-t}(v)\| \leq Ke^{-\lambda t}\|v\| \quad \text{for } v \in E^u,\ t \geq 0$$

($\lambda > 0$, K constants).

A *basic set* Ω of an Axiom A flow ϕ_t is a hyperbolic set in which periodic orbits are dense, $\phi_t|\Omega$ is topologically transitive and $\Omega = \bigcap_{t \in \mathbb{R}} \phi_t U$ for some open neighbourhood of Ω. *We shall always take a basic set Ω to be non-trivial*: *i.e.*, *Ω is not a topological circle.*

We shall be concerned with the restriction of an Axiom A flow ϕ_t to a basic set Ω so that ϕ_t shall stand for $\phi_t|\Omega$.

Bowen has shown that there are two mutually exclusive types of flow ϕ in this category: Either ϕ is a constant suspension of a homeomorphism (i.e. ϕ

factors over a circle) or the stable manifold of each point in Ω is dense in Ω. Moreover there is a unique ϕ invariant measure m on Ω giving maximum entropy to ϕ_1 [4]. If ϕ does not factor over a circle then ϕ is mixing with respect to m [4]. In other words:

PROPOSITION 1. *If ϕ is the restriction of an Axiom A flow to a basic set Ω, then either*

(i) *ϕ has a non-constant continuous eigenfunction*

$$g(\phi_t x) = e^{iat} g(x), \qquad g \in C(\Omega), \quad a \neq 0$$

or

(ii) *ϕ is mixing with respect to the maximal measure m.*

These possibilities are mutually exclusive.

2. Shifts of finite type

Let A be a zero-one irreducible $k \times k$ matrix and let

$$\Sigma_A = \left\{ x \in \prod_{n=-\infty}^{\infty} \{1, 2, \ldots, k\} : A(x_n, x_{n+1}) = 1, n \in \mathbb{Z} \right\}.$$

Then Σ_A is a compact zero-dimensional σ invariant set where the shift σ is defined by $(\sigma x)_n = x_{n+1}$. The *shift of finite type* defined by A is $\sigma_A = \sigma | \Sigma_A$.

If f is a strictly positive continuous function defined on Σ_A we define Σ_A^f as the set

$$\{(x, t) : 0 \leq t \leq f(x), x \in \Sigma_A\}$$

with $(x, f(x))$ and $(\sigma_A x, 0)$ identified. The f *suspension* σ^f of σ_A is the vertical flow defined on Σ_A^f by the local flow $\sigma_t^f(x, s) = (x, s + t)$ when $0 \leq s \leq f(x)$, $0 \leq s + t \leq f(x)$.

For $f \in C(\Sigma_A)$ and $0 < \theta < 1$ we define

$$\mathrm{var}_n f = \sup\{|f(x) - f(y)| : x, y \in \Sigma_A, x_i = y_i, |i| \leq n\}$$

and

$$\|f\|_\theta = \sup\{\mathrm{var}_n f / \theta^{2n+1} : n = 0, 1, \ldots\}.$$

The space $\mathscr{F}_\theta = \{f \in C(\Sigma_A) : \|f\|_\theta < \infty\}$ is a Banach space with respect to the norm $|||\ |||_\theta$ given by $|||f|||_\theta = \max(\|f\|_\infty, \|f\|_\theta)$.

For $f \in C(\Sigma_A)$ the *pressure* is defined as

$$\mathscr{P}(f) = \sup\left\{ \int f \, d\mu + h_\mu(\sigma_A) : \mu \ \sigma_A\text{-invariant} \right\},$$

and when $\mathscr{P}(f) = \int f \, d\mu + h_\mu(\sigma_A)$, μ is called an *equilibrium state* for f. Equilibrium states always exist in this setting.

Moreover [12], when $f \in \mathscr{F}_\theta$ there is exactly one equilibrium state for f. If $f \in \mathscr{F}_\theta$ and $f > 0$ then there exists $c > 0$ such that $\mathscr{P}(-cf) = 0$. Hence, there exists q such that

$$h_q(\sigma_A) - c\int f\,dq = 0 \qquad \text{and}$$

$$h_\mu(\sigma_A) - c\int f\,d\mu < 0$$

for all other σ_A invariant probabilities μ. Thus, using a result of Abramov's [1], one has $c = h(\sigma_1^f) = h(\sigma^f)$ the *topological entropy* of the f suspension.

Proposition 2. *When $f \in \mathscr{F}_\theta$, $f > 0$, we have $\mathscr{P}(-h(\sigma^f)f) = 0$.*

The probability measure p defined locally by the direct product of q with Lebesgue measure is the unique σ^f invariant probability giving topological entropy

$$h_p(\sigma_1^f) = h(\sigma_1^f) = h(\sigma^f).$$

3. Axiom A flows and suspensions

Drawing an analogy with his result for Axiom A diffeomorphisms, Bowen [5] proved the following powerful result.

Proposition 3. *If ϕ_t is the restriction of an Axiom A flow to a basic set Ω then there exist a shift of finite type σ_A, a positive function $f \in \mathscr{F}_\theta$ (for some $0 < \theta < 1$) and a continuous surjective map π: $\Sigma_A^f \to \Omega$ such that $\phi_t\pi = \pi\sigma_t^f$. Moreover there exists N such that* card $\pi^{-1}(y) \le N$ *for all $y \in \Omega$. π is measure-preserving with respect to the maximal entropy measures and with respect to these measures it is almost everywhere one-one.*

We refer to σ^f as the *principal suspension*.

Corollary 1. *ϕ is topologically weak-mixing if and only if σ^f is topologically weak-mixing.*

If ϕ is not topologically weak-mixing then certainly σ^f is not topologically weak-mixing. If ϕ is topologically weak-mixing then by Proposition 1, ϕ is measure-theoretically mixing with respect to the maximal measure. Hence σ^f is measure-theoretically mixing with respect to its maximal measure, so that it must be topologically weak-mixing.

Corollary 2. $h(\phi) = h(\sigma^f)$.

This follows from the fact that π is at most N to 1, or from the fact that π is measure-preserving and almost everywhere one-one with respect to maximal measures.

Our aim is to count the number of closed orbits of the flow ϕ. To do this Bowen [5] introduced, following Manning [8] for the diffeomorphism case, auxiliary shift suspensions.

PROPOSITION 4. *Alongside the principal suspension $\sigma^f = \sigma^0$, there exist suspensions $\sigma^1, \ldots, \sigma^m$ of other shifts of finite type with the property that*

$$\nu(\phi, x) = \sum_{i=1}^{m} (-1)^{l(i)} \nu(\sigma^i, x) + \nu(\sigma^0, x) \tag{3.1}$$

where, for example, $\nu(\phi, x)$ is the number of closed orbits of ϕ with least period x. Moreover if $i \neq 0$ then there exists a continuous map π_i with range properly contained in Ω such that π_i is at most N to 1 and $\pi_i \sigma_t^i = \phi_t \pi_i$.

Remark. Since each π_i ($i \neq 0$) is at most N to 1 with range a proper closed subset of Ω, it follows that the topological entropies of σ^i ($i \neq 0$) are strictly less than $h(\sigma^0) = h(\sigma^f) = h(\phi)$.

4. The zeta function

If ψ is a flow on a compact space we denote by τ a generic closed orbit and by $\lambda(\tau)$ the least period of that orbit. The zeta function of the flow is defined by

$$\zeta(s) = \prod_{\tau} \frac{1}{(1 - N(\tau)^{-s})} \tag{4.1}$$

where this makes sense. Here, $N(\tau) = e^{\lambda(\tau)}$ is the *norm* of τ.

As long as $\sum_{\tau} \sum_{k=1}^{\infty} \frac{1}{k} N(\tau)^{-sk}$ converges absolutely, (4.1) is well defined and

$$\zeta(s) = \exp \sum_{\tau} \sum_{k=1}^{\infty} \frac{1}{k} N(\tau)^{-sk}. \tag{4.2}$$

If σ^f is the f suspension of the shift of finite type σ_A, then

$$\sum_{\tau} \sum_{k=1}^{\infty} \frac{1}{k} N(\tau)^{-sk} = \sum_{k=1}^{\infty} \frac{1}{k} \sum_{x \in \operatorname{fix} \sigma_A^k} \exp - s\left(f(x) + \cdots + f(\sigma_A^{k-1} x)\right)$$

which converges absolutely when

$$\lim_{k \to \infty} \left(\sum_{x \in \operatorname{fix} \sigma_A^k} \exp - \sigma\left(f(x) + \cdots + f(\sigma_A^{k-1} x)\right)\right)^{1/k} = e^{\mathscr{P}(-\sigma f)} < 1,$$

where $s = \sigma + it$ (cf. [12]). In view of Proposition 2 and the monotonicity of $\mathscr{P}$, therefore, we have

PROPOSITION 5. *For the f suspension of a shift of finite type, $\zeta(s)$ is well defined, analytic and non-zero when $\mathscr{R}(s) > h(\sigma^f)$.*

Let ϕ be the restriction of an Axiom A flow to a basic set, and let σ^i, $i = 0,\dots,m$ be the suspensions of shifts of finite type mentioned in Proposition 4 (σ^0 is the principal suspension and σ^i, $i \neq 0$, are auxiliary suspensions). If we combine (3.1) and (4.2) we see that

$$\zeta(s) = \zeta_0(s)\prod_{i=1}^{m}(\zeta_i(s))^{(-1)^{l(i)}}$$

where ζ is the zeta function of ϕ and ζ_i ($i = 0,\dots,m$) is the zeta function of σ^i. Since ϕ and σ^0 have the same topological entropy h and σ^i, $i = 1,\dots,m$, have topological entropies less than h we see the following:

PROPOSITION 6. *$\zeta(s) = \zeta_0(s)\cdot\eta(s)$ for $\mathscr{R}(s) > h$ where $\eta(s)$ is analytic and non-zero in $\mathscr{R}(s) > h - \varepsilon$ for some $\varepsilon > 0$ ($h - \varepsilon$ can be taken as the maximum of the entropies of σ^i, $i = 1,\dots,m$).*

In the next few sections we shall meromorphically extend ζ, ζ_0 to an open domain U containing $\mathscr{R}(s) \geq h$. In general, these extensions will be analytic in $U - \{h\}$ with a simple pole at h. We shall, however, encounter an exceptional case where ζ, ζ_0 are periodic with period it_0 ($t_0 > 0$) and where their extensions have simple poles at $h + nit_0$, $n \in \mathbb{Z}$. This case corresponds exactly to the condition that ϕ, σ^0 are *not* topologically weak mixing. In view of Proposition 6 the problem of extending ζ reduces to the problem of extending ζ_0.

5. Locally constant functions

In this section we shall consider the special case where our strictly positive function f, defined on (Σ_A, σ_A) is locally constant; i.e., f assumes only a finite number of values. In other words there is some $n > 0$ such that f is constant on each cylinder

$$[i_{-n},\dots,i_0,\dots,i_n]_{-n} = \{x \in \Sigma_A : x_j = i_j \quad \text{for } |j| \leq n\}.$$

By a simple coding (regarding words of length $2n$ as new symbols) there is no loss in generality in assuming f to be a function of two coordinates $f(x) = f(x_0, x_1)$.

PROPOSITION 7 [10]. *The flow σ^f is not weak-mixing with eigenfrequency $a > 0$ if and only if there exists an integer-valued continuous function M and a* *continuous function u assuming only finitely many values such that* $af = 2\pi M + u \circ \sigma_A - u$.

From this we deduce that if σ^f is not weak-mixing then $\exp(f(x_0, x_1) + \cdots + f(x_{n-1}, x_0))$ is an integer power of $\exp(2\pi/a)$. Conversely if $\exp(f(x_0, x_1) + \cdots + f(x_{n-1}, x_0))$ is an integer power of a single number for each closed orbit then it is relatively easy to show that there exists an integer-valued function M of two variables and $a > 0$ such that

$$af(i, j) = 2\pi M(i, j) + u(j) - u(i)$$

for some u. In other words:

Proposition 8. *The flow σ^f is not weak-mixing if and only if the multiplicative group $\Gamma_f = \langle \exp(f(x_0, x_1) + \cdots + f(x_{n-1}, x_0))\rangle$ generated by numbers $N(\tau)$ has rank* 1.

The Perron-Frobenius theorem assures us that the matrix $\{A(i, j) \cdot \exp - hf(i, j)\}$ has a maximum positive eigenvalue $\exp \mathscr{P}(-hf) = 1$ with an associated positive eigenvector r such that

$$\sum_j A(i, j) r_j \exp - hf(i, j) = r_i .$$

Hence the matrix $P(i, j) = \left\{A(i, j)\frac{r_j}{r_i}\exp - hf(i, j)\right\}$ is stochastic. For simplicity we shall assume that $h = 1$.

The zeta function of σ^f is

$$\zeta(s) = \exp \sum_{n=1}^{\infty} \frac{1}{n} \sum_{x \in \mathrm{fix}\, \sigma_A^n} \exp - s\left(f(x) + \cdots + f(\sigma_A^{n-1} x)\right)$$

$$= \exp \sum_{n=1}^{\infty} \frac{1}{n} \mathrm{Trace}(P^s)^n$$

where P^s is the matrix with complex entries $P(i, j)^s$. Thus $\zeta(s) = \dfrac{1}{\det(I - P^s)}$ for $\mathscr{R}(s) > 1$ and this closed form for $\zeta(s)$ provides a non-vanishing meromorphic extension to $\mathbf{C}$.

Since $s = 1$ is a simple eigenvalue of P there is a simple pole at $s = 1$. To see this let $\beta_1(s), \beta_2(s), \ldots$ be eigenvalues of P^s in a small neighbourhood of $s = 1$ and write $\det(I - P^s) = \prod_{j=1}^{k}(1 - \beta_j(s))$. We may assume that $\beta_1(1) = 1$, that $\beta_1(s)$ is analytic and $|1 - \beta_j(s)| > \varepsilon > 0$ for $j \neq 1$ and all s in a small neighbourhood of $s = 1$. Hence if $\det(I - P^s) = (s-1)^2\phi(s)$ where ϕ is analytic then

$$|(s-1)\phi(s)| \geq \varepsilon^{k-1}\left|\frac{1-\beta_1(s)}{s-1}\right|$$

so that $\beta_1'(1) = 0$. However one can show (cf. [11]) that $-\beta_1'(1)$ is the entropy of σ_A, with respect to the Markov measure defined by P, which is certainly not zero. We deduce, therefore, that the zero of $\det(I - P^s)$ at 1 is simple.

Suppose $\zeta(s)$ has a pole elsewhere on $\mathscr{R}(s) = 1$; i.e., $\det(I - P^s)$ has a zero at $s = 1 + it_0$ with $|t_0|$ least. Then

$$\sum_k P(j,k)\cdot P(j,k)^{it_0}\xi_k = \xi_j \tag{5.1}$$

for some non-zero vector ξ and $\sum_k P(j,k)|\xi_k| \geq |\xi_j|$ so that $|\xi_k|$ is independent of k. We may suppose that $|\xi_j| = 1$ and (5.1) says that a convex combination of points on the unit circle is a point on the unit circle. This means that $P(j,k)^{it_0}\xi_k = \xi_j$ whenever $P(j,k) \neq 0$. We see that $P(x_0,x_1)^{-it_0}\cdots P(x_{n-1},x_0)^{-it_0} = 1 = \exp it_0(f(x_0,x_1) + \cdots + f(x_{n-1},x_0))$ for all closed orbits. In other words $\exp(f(x_0,x_1) + \cdots + f(x_{n-1},x_0)) = \alpha^m$ for some m where $\alpha = \exp\dfrac{2\pi}{t_0}$. From Proposition 8 we have that σ^f is not weak-mixing and therefore:

PROPOSITION 9. *If f is positive and locally constant then ζ has a nowhere vanishing meromorphic extension to* $\mathbf{C}$.

If σ^f is not topologically weak-mixing then there exists a least $t_0 > 0$ such that ζ is simply periodic with period it_0. ζ is analytic in $\mathscr{R}(s) > h - \varepsilon$, for some $\varepsilon > 0$, except for simple poles at $h + nit_0$, $n \in \mathbf{Z}$.

If σ^f is topologically weak-mixing then there exists an open domain U containing $\mathscr{R}(s) \geq h$ in which ζ is analytic except for the simple pole at $s = h$.

6. When ϕ is not topologically weak-mixing

In this section ϕ denotes the Axiom A flow restricted to Ω and $\sigma^0 = \sigma^f$ denotes the principal suspension of Section 3.

In this case ϕ, σ^0 have a common minimum positive eigenfrequency a:

$$g\phi_t = e^{iat}g, \qquad g \in C(\Omega), \quad |g| = 1;$$

$$g\pi_0\sigma_t^0 = e^{iat}g\pi_0, \qquad g\pi_0 \in C(\Sigma_A^f).$$

For the flow ϕ the minimum eigenfrequency a is evidently greater than or equal to the minimum eigenfrequency for σ^0. Since π_0 is a measure isomorphism between Ω and Σ_A^f with respect to their maximal measures, equality follows from the spectral equivalence of the flows.

We see, then, that the highest common factor of $\{\lambda(\tau): \tau$ a closed ϕ orbit$\}$ is a multiple of $\frac{2\pi}{a}$, say $\frac{2\pi m}{a}$. In the same way the highest common factor of $\{\lambda(\tau): \tau$ is a closed σ^0 orbit$\}$ is $\frac{2\pi n}{a}$, $n \in \mathbf{N}$. Since π_0 is finite to one and surjective, each σ^0 closed orbit τ_0 maps to a ϕ closed orbit τ with $\lambda(\tau_0)$ a multiple of $\lambda(\tau)$ so that n is a multiple of m.

From Section 1, σ^0 may be interpreted as a constant suspension and so there is no loss in generality in assuming that $n = 1$ from which we conclude that $m = n = 1$. From the equations

$$\zeta(s) = \exp \sum_{\tau} \sum_{k=1}^{\infty} \frac{1}{k} e^{-s\lambda(\tau)},$$

$$\zeta_0(s) = \exp \sum_{\tau} \sum_{k=1}^{\infty} \frac{1}{k} e^{-s\lambda(\tau)},$$

where in each case the sum is over the closed orbits of ϕ, σ^0 respectively, it follows that ζ, ζ_0 are simply periodic with period ia and moreover this is the *least* period.

Remark. When σ^0 is not topologically weak-mixing we can take f to be constant. Hence Proposition 9 applies and together with Proposition 6 we see that ζ, ζ_0 have non-zero analytic and simply periodic extensions to $\mathscr{R}(s) > h - \varepsilon$ except for the simple poles at $h + nia$, $n \in \mathbf{Z}$.

7. Ruelle's theorem

In this section we show, following Ruelle [12] (Exercise 7(c), pp. 100–101), how the zeta function of a suspension σ^f of σ_A may be extended to a neighbourhood of $s = h$ when σ^f is topologically weak-mixing.

Proposition 10 (Ruelle [12], p. 93). *If $f \in \mathscr{F}_\theta$ then there exists a continuous function $R: \mathscr{F}_\theta \to \mathbf{R}^+$ such that $R(f) > \exp - \mathscr{P}(f)$ and*

$$d(z, f) = \exp - \sum_{m=1}^{\infty} \frac{z^m}{m} \sum_{x \in \operatorname{Fix} \sigma_A^m} e^{f(x) + \cdots + f(\sigma_A^{m-1} x)}$$

extends to an analytic function in $\{z: |z| < R(f)\}$ with only one zero, which is simple and located at $\exp - \mathscr{P}(f)$.

The functions d and ζ are related by $\zeta(s) = \dfrac{1}{d(1, -sf)}$ so we need to extend this result to certain complex functions.

Consider a conformal mapping $g(z)$ of the unit disc to the upper semi-disc of radius 2 sending 0 to i. Define $G(z) = \frac{d^{(n)}}{n!}(0, -(\sigma + g(z)t)\cdot f)$ for σ, t real. ($d^{(n)}$ denotes the n'th derivative with respect to the first variable.) By Jensen's inequality,

$$|G(0)| \le \exp \frac{1}{2\pi}\int_0^{2\pi} \log|G(e^{i\theta})|\,d\theta$$

and so

$$\overline{\lim}\left|\frac{d^{(n)}}{n!}(0, -(\sigma + it)f)\right|^{1/n}$$

$$\le \exp\frac{1}{2\pi}\int_0^{2\pi}\log\overline{\lim}\left|\frac{d^{(n)}}{n!}(0, -(\sigma + g(e^{i\theta})t)f)\right|^{1/n} d\theta$$

$$\le \exp\frac{1}{2\pi}\int_0^{2\pi} -\log R(-(\sigma + g(e^{i\theta})t)f)\,d\theta, \tag{7.1}$$

where R is the function of Proposition 10 extended to equal $\exp -\mathscr{P}(\mathscr{R}(f))$ for f complex and not real valued. The application of Fatou's lemma, here, can be justified by finding a dominating function with the aid of Cauchy's formulae.

Since (7.1) is continuous with respect to $\sigma + it$ and strictly less than $\exp \mathscr{P}(-\sigma f)$ when $t = 0$, we see that

$$\frac{1}{\overline{\lim}\left|\frac{d^{(n)}}{n!}(0, -(\sigma + it)f)\right|^{1/n}} > \exp -\mathscr{P}(-hf) = 1$$

in a sufficiently small neighbourhood U of $s = h$. Thus $s \to d(1, -sf)$ is analytic on U and by studying Ruelle's extension we see, as for the locally constant case, $s = h$ is a simple zero.

In fact the proof of Proposition 10 reveals that the essential part of the function $d(z, -\sigma f)^{-1}$ (σ real) is $\dfrac{C}{(1 - ze^{\mathscr{P}(-\sigma f)})}$ $(C \neq 0)$ in $|z| < R(-\sigma f)$ so that the essential part of $\zeta(\sigma)$ is $\dfrac{C}{(1 - e^{\mathscr{P}(-\sigma f)})}$, and the proof that $s = h$ is a simple pole of $\zeta(s)$ follows along the lines of the proof in the locally constant case. One need only check that $e^{\mathscr{P}(-\sigma f)}$ has a non-zero derivative (with respect to σ) at $\sigma = h$. In fact the derivative is $-\int f\,dm$, where m is the equilibrium state for $-hf$. We therefore have

Proposition 11 [12]. *Let $f \in \mathscr{F}_\theta$. Then ζ has an analytic extension to a neighbourhood of $s = h$ except for a simple pole at $s = h$.*

8. Extensions beyond $\mathscr{R}(s) = h$

As usual we suppose that f is a positive function belonging to $\mathscr{F}_\theta$ for some $0 < \theta < 1$. It is known (cf. [6]) that such a function is *cohomologous* to a "function of the future" f' belonging to $\mathscr{F}_{\theta^{1/2}}$; i.e., $f'(x) = f'(y)$ if $x_i = y_i$, $i = 0, 1, \ldots$, and there exists a continuous u with

$$f + u \circ \sigma_A - u = f'. \tag{8.1}$$

Moreover, since f is bounded away from zero, $\dfrac{1}{n}(f' + f' \circ \sigma_A + \cdots + f' \circ \sigma_A^{n-1})$ is strictly positive for large enough n, and the latter is cohomologous to f'. Hence, in (8.1), f' may be taken as a strictly positive function of the future. It is clear that the zeta functions associated with two functions coincide if they differ by a *coboundary* (a function of the form $u \circ \sigma_A - u$) so there will be no loss in generality if we assume, as we shall, that f is a strictly positive function of the future, belonging to $\mathscr{F}_{\theta^{1/2}}$. As such, f can be thought of as a function defined on

$$\Sigma_A^+ = \left\{ x \in \prod_{n=0}^{\infty} \{1, 2, \ldots, k\} : A(x_n, x_{n+1}) = 1 \right\}.$$

Associated with Σ_A^+ we have the *one-sided shift of finite type*, which we denote by σ_A. (We hope this will not lead to confusion.) Also associated with Σ_A^+ we have the spaces

$$\mathscr{F}_\theta^+ = \{ k \in C(\Sigma_A^+) : \sup(\mathrm{var}_n k/\theta^n) < \infty \} \quad \text{and}$$

$$\mathscr{F}_\theta^{\mathrm{C}} = \{ k \in C(\Sigma_A^+, \mathrm{C}) : \sup(\mathrm{var}_n k/\theta^n) < \infty \}.$$

These are Banach spaces with respect to the norm $|||k|||_\theta = \max(\|k\|_\theta, \|k\|_\infty)$, where $\|k\|_\theta = \sup(\mathrm{var}_n k/\theta^n)$.

If $k \in \mathscr{F}_\theta^{\mathrm{C}}$, the *Ruelle operator* $\mathscr{L}_k : \mathscr{F}_\theta^{\mathrm{C}} \to \mathscr{F}_\theta^{\mathrm{C}}$ is defined by

$$\sum_{\sigma_A y = x} e^{k(y)} v(y) = (\mathscr{L}_k v)(x).$$

Since f may be viewed as a function defined on Σ_A^+ we have $f \in \mathscr{F}_\theta^+$. We shall be especially interested in the operators $\mathscr{L}_{sf} : \mathscr{F}_\theta^{\mathrm{C}} \to \mathscr{F}_\theta^{\mathrm{C}}$, s complex. There is one more legitimate and convenient assumption concerning f which we wish to make. The Ruelle-Perron-Frobenius theorem (cf. [12]) guarantees the existence of $u > 0$, $u \in \mathscr{F}_\theta^+$ such that

$$\mathscr{L}_{-hf} u = u$$

so that, by replacing $-hf$ by $-hf + \log u - \log u \circ \sigma_A$ if necessary, we can assume $\mathscr{L}_{-hf} 1 = 1$. (This equation ensures that we still have $f \geq 0$, although it is conceivable that f may not be strictly positive.)

Fix $t \in \mathbf{R} - \{0\}$ and for the remainder of this section define $L = \mathscr{L}_{-(h+it)f}$. There are two cases to be considered corresponding to whether or not

$$(8.2) \qquad Lg = e^{i\alpha}g \quad \text{for some } 0 \leq \alpha < 2\pi, \text{ and non-trivial } g \in \mathscr{F}_\theta^{\mathbf{C}}.$$

(i) Assume that (8.2) has a non-trivial solution then by the argument of Section 5,

$$e^{-i(tf(x)+\alpha)}g(x) = g(\sigma_A x)$$

and therefore, writing $g(x) = e^{ir(x)}$ with $r\colon \Sigma_A \to \mathbf{R}$ continuous, we have

$$-(tf+\alpha) + r - r \circ \sigma_A = 2\pi M$$

where $M\colon \Sigma_A \to \mathbf{R}$ is integer valued and continuous. Hence f is cohomologous to a locally constant function. The zeta function is therefore described in Section 5, Proposition 9.

Proposition 12. *If* (8.2) *has a non-trivial solution then ζ has a meromorphic extension to* C. *If Γ_f is of rank* 1 *then ζ, the extended function, is simply periodic (period it_0), nowhere vanishing and analytic in $\mathscr{R}(s) > h - \varepsilon$ ($\varepsilon > 0$) except for simple poles at $h + nit_0$, $n \in \mathbf{Z}$. If Γ_f is not of rank* 1 *(rank* 2 *in fact) then ζ is nowhere vanishing and analytic in an open neighbourhood of $\mathscr{R}(s) \geq h$ except for a simple pole at $s = h$.*

(ii) In this case we assume that (8.2) has no non-trivial solution. In due course we shall prove that $\rho(L) < 1$ where ρ is the *spectral radius*. To this end we prove the following:

Proposition 13. *If* (8.2) *has no non-trivial solution then for every $k > 0$ there exists N such that for all $n > N$*

$$\sup\{\|L^n g\|_\infty : \|g\|_\infty \leq 1, |||g|||_\theta \leq k\} < 1.$$

Proof. If this is not the case then for some $k > 0$ there exist infinitely many $n \in \mathbf{N}$ such that

$$\sup\{\|L^n g\|_\infty : \|g\|_\infty \leq 1, |||g|||_\theta \leq k\} = 1$$

and since $\{g\colon \|g\|_\infty \leq 1, |||g|||_\theta \leq k\}$ is a $\|\ \|_\infty$ equicontinuous and therefore compact set in $\{g\colon \|g\|_\infty \leq 1\}$, we have $\|L^n g_n\|_\infty = 1$ for infinitely many n where $|||g_n|||_\theta \leq k$. Let g be an accumulation point of $\{g_n\}$ with respect to the norm $\|\ \|_\infty$. Since $\|L^k g_n\|_\infty = 1$ for $k \leq n$ (L being a $\|\ \|_\infty$ contraction), we have $\|L^k g\|_\infty = 1$ for $k \geq 0$.

For each $k \geq 0$ choose x^k so that

$$|(L^k g)(x^k)| = 1 = \left| \sum_{\sigma_A^k y = x^k} e^{-hf^k(y)} e^{-itf^k(y)} g(y) \right|$$

where $f^k(y) = f(y) + \cdots + f(\sigma_A^{k-1} y)$. The standard convexity argument (used in §5 and in (i) above) shows that $e^{-itf^k(y)} \cdot g(y)$ and $|g(y)|$ are independent of the choice of y such that $\sigma_A^k y = x^k$. Since $\bigcup_{k=0}^{\infty} \{ y \colon \sigma_A^k y = x^k \}$ is dense in Σ_A we conclude that $|g| = 1$. However, the same argument applies to $L^n g$ and therefore $|L^n g| = 1$ for all $n \geq 0$. Now fix n and x so that $\sigma_A^n x = x$. Clearly

$$(L^{kn} g)(x) = \sum_{\sigma_A^{kn} y = x} e^{-hf^{kn}(y)} \cdot e^{-itf^{kn}(y)} g(y)$$

and again by the convexity argument above we have

$$e^{-itf^{kn}(y)} g(y) = e^{-itf^{kn}(x)} g(x) \tag{8.3}$$

whenever $\sigma_A^{kn} y = x$. Comparing (8.3) with a similar equation for Lg we have $(Lg)(y)/g(y)$ is constant for all y such that $\sigma_A^{kn} y = x$, for all $k \geq 0$. When A is aperiodic we deduce that we have a non-trivial solution of (8.2) contradicting our assumption. When A is not aperiodic a similar but slightly more complicated argument can be given. The proposition is proved.

We shall need the inequality

$$\|L^k g\|_\theta \leq C\|g\|_\infty + \|g\|_\theta \theta^k \tag{8.4}$$

(some constant $C > 0$), which is easily established, in the proof of

Proposition 14. *If $Lg = e^{ia} g$ has no non-trivial solution then $\rho(L) < 1$.*

Proof. Choose g with $|||g|||_\theta \leq 1$; then for all $n \geq 0$, $L^n g \in \{ k \colon |||k|||_\theta \leq C + 1 \}$ and by the $\| \ \|_\infty$ compactness of this ball there is an accumulation point k of $\{ L^n g \}$. Hence

$$\|g\|_\infty \geq \|Lg\|_\infty \geq \cdots \geq \|k\|_\infty = \|L^n k\|_\infty$$

for all $n \geq 0$. If $\|k\|_\infty \neq 0$ then $k/\|k\|_\infty$ contradicts Proposition 13. Consequently $\|L^n g\|_\infty \to 0$ as $n \to \infty$ and we have, by (8.4),

$$\|L^{2n} g\|_\theta \leq C\|L^n g\|_\infty + \theta^n [C \cdot \|g\|_\infty + \|g\|_\theta \theta^n] < 1$$

for large n. By the $\| \ \|_\infty$ compactness of $\{ k \colon |||k|||_\theta \leq 1 \}$, we have

$$\sup\{ \|L^n g\|_\theta \colon |||g|||_\theta \leq 1 \} < 1$$

for large n. Hence

$$\rho(L) = \inf \sup\{|||L^n g|||_\theta^{1/n} : |||g|||_\theta \le 1\} < 1.$$

Assume wlog $f \in \mathcal{F}_{\theta_0}{}^{n}$ $(\theta_0 < \Theta)$

We continue with the assumption that (8.2) has no non-trivial solution.

Choose $0 < \alpha < 1$ such that $k^\alpha \rho(L) < 1$ (where A is $k \times k$) and write L^s for $\mathscr{L}_{-sf}$. Because the maps $s \to \theta^\alpha e^{\mathscr{P}(-\sigma f)}$, $s \to \rho(L^s) \cdot k^\alpha$ (where $s = \sigma + iv$) are continuous and upper semi-continuous respectively and both are strictly smaller than 1 at $s = h + it$, we have

$$\theta^\alpha e^{\mathscr{P}(-\sigma f)}, \rho(L^s) \cdot k^\alpha < 1 \quad \text{for } s = \sigma + iv \in D$$

where $D = D(h + it)$ is a sufficiently small disc with centre $h + it$.

Define a locally constant function f_n by $f_n(x) = \sup\{f(z): \ z_i = x_i, 0 \le i \le n\}$; then $\|f_n - f\|_\infty \le \|f\|_\theta \cdot \theta^n$ and $|||f_n - f|||_\Theta \le |f|_{\theta_0} (\theta_0/\Theta)^n$

If $n, m \ge 0$ then

$$\zeta(s) = \exp\left(\sum_{m=1}^\infty \frac{1}{m} \sum_{x \in \mathrm{fix}\,\sigma_A^m} \left(e^{-sf^m(x)} - e^{-sf_n^m(x)}\right)\right)$$

$$\times \exp\left(\sum_{m=1}^\infty \frac{1}{m} \sum_{x \in \mathrm{fix}\,\sigma_A^m} e^{-sf_n^m(x)}\right).$$

Then putting $s = \sigma + iv \in D$ and $n = [m\alpha]$ we have

$$\overline{\lim}\left|\sum_{x \in \mathrm{fix}\,\sigma_A^m} \left(e^{-sf^m(x)} - e^{-sf_n^m(x)}\right)\right|^{1/m}$$

$$\le \overline{\lim}\left|\sum_{x \in \mathrm{fix}\,\sigma_A^m} \left(e^{-\sigma f^m(x)}|e^{-ivf^m(x)} - e^{-ivf_n^m(x)}|\right.\right.$$

$$\left.\left. + |e^{-\sigma f^m(x)} - e^{-\sigma f_n^m(x)}|\,|e^{-ivf_n^m(x)}|\right)\right|^{1/m}$$

$$\le \overline{\lim}\left|\sum_{x \in \mathrm{fix}\,\sigma_A^m} e^{-\sigma f^m(x)} \cdot m \cdot \|f\|_\theta \theta^n \left(|v| + |\sigma| \cdot e^{m\|f\|_\theta \theta^n}\right)\right|^{1/m}$$

$$\le e^{\mathscr{P}(-\sigma f)} \theta^\alpha < 1.$$

Notice that $\Sigma_{x \in \mathrm{fix}\,\sigma_A^m} e^{-sf_n^m(x)} = \mathrm{Trace}(P_n^s)^m$ where P_n^s is the operator induced by $\mathscr{L}_{-sf_n}$ on the finite dimensional subspace of $C(\Sigma_A^+, \mathbf{C})$ with base vectors

$$\delta_{[z_0, \ldots, z_{n-1}]}(x) = 1 \quad \text{if } x_i = z_i, \qquad 0 \le i < n$$
$$= 0 \qquad \text{otherwise.}$$

The eigenvalues of P_n^s, $(\lambda_1, \ldots, \lambda_{N(m)})$, are contained in the spectrum of $\mathscr{L}_{-sf_n}$

and $N(m) \leq k^n$. Hence

$$\overline{\lim}\left|\sum_{x \in fix\,\sigma_A^m} e^{-sf_n^m(x)}\right|^{1/m} = \overline{\lim}\left|\lambda_1^m + \cdots + \lambda_{N(m)}^m\right|^{1/m}$$

$$\leq \rho(L^s) \cdot k^a < 1.$$

Thus ζ is analytic and non-zero on the disc D. Combining this with Ruelle's theorem (Proposition 11) we have

Proposition 15. *If $Lg = e^{ia}g$ has no non-trivial solution then ζ has a non-zero analytic extension to a neighbourhood of $\mathscr{R}(s) \geq h$, except for a simple pole at $s = h$. (In particular the flow σ^f is topologically weak-mixing.)*

9. Extending ζ for an Axiom A flow

Combining the remark of Section 6 and Propositions 12 and 15, we have now proved, for ϕ an Axiom A flow restricted to a basic set, the following:

Theorem 1. *If ϕ is not topologically weak-mixing then ζ has a nowhere vanishing extension to $\mathscr{R}(s) > h - \varepsilon$ (some $\varepsilon > 0$) in which ζ is simply periodic with least period ia (where $a > 0$ is the least eigenfrequency) and ζ is analytic except for simple poles at $h + nia$, $n \in \mathbb{Z}$.*

If ϕ is topologically weak-mixing then ζ has a nowhere vanishing analytic extension to an open neighbourhood of $\mathscr{R}(s) \geq h$ except for a simple pole at $s = h$.

10. Number theory

In this section we count the number of closed orbits of the Axiom A flow ϕ restricted to a basic set and obtain asymptotic formulas by following the Wiener-Ikehara [15] proof of the prime number theorem. The ζ function for the flow will play the role of Riemann's zeta function in the proof. For each closed orbit τ of ϕ define $N_h(\tau) = e^{\lambda(\tau)h}$ where $\lambda(\tau)$ is the least period of τ and h is the topological entropy of ϕ.

For purely computational purposes it is convenient to introduce fictitious orbits and an analogue of von Mangoldt's function. A fictitious orbit is a formal product

$$\tau' = \tau_1^{l_1} \cdots \tau_m^{l_m}$$

where τ_i are genuine closed orbits and l_i are positive integers. (Primed orbits will always denote fictitious orbits. Genuine orbits will have no prime.) For such an object define

$$N_h(\tau') = N_h(\tau_1)^{l_1} \cdots N_h(\tau_m)^{l_m}$$

and

$$\Lambda(\tau') = \log N_h(\tau) \quad \text{if } \tau' = \tau^l, l \geq 1$$
$$= 0 \qquad \text{otherwise.}$$

Define

$$\zeta_h(s) = \zeta(sh) = \prod_\tau \left(1 - N_h(\tau)^{-s}\right)^{-1}$$
$$= \exp \sum_\tau \sum_{k=1}^{\infty} \frac{1}{k} N_h(\tau)^{-ks}$$

which converge when $\mathscr{R}(s) > 1$. (The critical behaviour we investigated in Section 8 now occurs on the line $\mathscr{R}(s) = 1$.)

Taking logarithmic derivatives we have

$$\text{(10.1)} \qquad \frac{\zeta_h'(s)}{\zeta_h(s)} = -\sum_\tau \log N_h(\tau) \sum_{m=1}^{\infty} N_h(\tau)^{-ms}$$
$$= -\sum_{\tau'} \frac{\Lambda(\tau')}{N_h(\tau')^s} \quad \text{provided } \mathscr{R}(s) > 1.$$

We consider the two cases when ϕ is or is not weak-mixing separately.

1) If ϕ is not weak-mixing then there exists $\alpha = \exp \dfrac{2\pi h}{a}$ such that $N_h(\tau')$ is a positive integral power of α (a is the least eigenfrequency). Thus

$$\text{(10.2)} \qquad \frac{\zeta_h'(s)}{\zeta_h(s)} = -\sum_{n=1}^{\infty} \alpha^{-ns} \sum_{N_h(\tau') = \alpha^n} \Lambda(\tau').$$

Using simple periodicity and noting that $\zeta_h'(s)/\zeta_h(s)$ has simple poles at $1 + nia/h$, $n \in \mathbb{Z}$, with residue -1, we have

$$\frac{\zeta_h'(s)}{\zeta_h(s)} = \frac{-2\pi h/a}{1 - e^{-2\pi h(s-1)/a}} + \phi(s)$$

where $\phi(s)$ is analytic and simply periodic in $\mathscr{R}(s) > 1 - \varepsilon$.

Hence in the same region we have

$$\frac{\zeta_h'(s)}{\zeta_h(s)} = -\frac{2\pi h}{a} \sum_{n=0}^{\infty} e^{(2\pi nh)/a} \cdot e^{-(2\pi nhs)/a} + \phi(s)$$

and by (10.2) we conclude that

$$\sum_{N_h(\tau') = \alpha^n} \Lambda(\tau') - \frac{2\pi h}{a} \cdot e^{(2\pi nh)/a}$$

converges to zero exponentially fast (cf. [12]). Hence defining

$$F(y) = \sum_{N_h(\tau') \le y} \Lambda(\tau'),$$

we have

Proposition 16. *If ϕ is not topologically weak-mixing then*

$$F(y) \sim \log \alpha \sum_{\alpha^n \le y} \alpha^n .$$

2) If ϕ is topologically weak-mixing then by (10.1) $\dfrac{\zeta_h'(s)}{\zeta_h(s)} = -\displaystyle\int_1^\infty x^{-s}\, dF(x)$ where, as before, $F(x) = \sum_{N_h(\tau') \le x} \Lambda(\tau')$.

Evidently

$$\int_1^\infty x^{-s}\, dF(x) = \frac{1}{s-1} - \psi(s)$$

where ψ is analytic in a neighbourhood of $\mathscr{R}(s) \ge 1$. Ikehara's Tauberian theorem (cf. [15]) therefore ensures that

Proposition 17. *If ϕ is topologically weak-mixing $F(y) \sim y$.*

Whether or not ϕ is weak-mixing we have

Proposition 18. $F(x) \sim \pi(x) \cdot \log x$, *where* $\pi(x) = \{\tau : N_h(\tau) \le x\}$.

Proof. By definition,

$$\begin{aligned} F(x) &= \sum_{N_h(\tau') \le x} \Lambda(\tau') \\ &= \sum_{N_h(\tau)^k \le x} \log N_h(\tau) = \sum_{\substack{N_h(\tau)^k \le x \\ k \text{ highest}}} k \log N_h(\tau). \end{aligned}$$

Since in the latter sum $k = \left[\dfrac{\log x}{\log N_h(\tau)}\right]$ we have

$$\begin{aligned} F(x) &\le \sum_{N_h(\tau) \le x} \left[\frac{\log x}{\log N_h(\tau)}\right] \cdot \log N_h(\tau) \\ &\le \sum_{N_h(\tau) \le x} \log x = \pi(x) \cdot \log x. \end{aligned}$$

For the other direction, we need to show first that $\dfrac{\pi(x)}{x^{\sigma}} \to 0$ as $x \to \infty$ whenever $\sigma > 1$:

$$\zeta(\sigma) = \prod_{\tau}\left(1 - N_h(\tau)^{-\sigma}\right)^{-1} \geq \prod_{N_h(\tau) \leq x} \left(1 - x^{-\sigma}\right)^{-1}$$

$$= \left(1 - x^{-\sigma}\right)^{-\pi(x)}.$$

Hence $\dfrac{\log \zeta(\sigma)}{x^{\sigma' - \sigma}} \geq \dfrac{\pi(x)}{x^{\sigma'}}$, $(\sigma' > \sigma)$, and therefore $\pi(x)/x^{\sigma'} \to 0$ as $x \to \infty$.

Now let $\sigma < 1$ and define $y = \left(\dfrac{x}{\log x}\right)^{\sigma}$. Evidently,

$$\begin{aligned}
\pi(x) &= \pi(y) + \sum_{y < N_h(\tau) \leq x} 1 \\
&\leq \pi(y) + \sum_{N_h(\tau) \leq x} \frac{\log N_h(\tau)}{\log y} \\
&\leq \pi(y) + \frac{1}{\log y} \sum_{N_h(\tau') \leq x} \Lambda(\tau') \\
&= \pi(y) + \frac{F(x)}{\log y}.
\end{aligned}$$

Hence

$$\pi(x) \leq \pi(y) + \frac{F(x)}{\sigma(\log x - \log\log x)}$$

and

$$\frac{F(x)}{x} \leq \frac{(\log x)\pi(x)}{x} \leq \frac{\pi(y)}{y^{1/\sigma}} + \frac{F(x)}{\sigma x} \cdot \frac{\log x}{(\log x - \log\log x)}.$$

Since $\dfrac{\pi(y)}{y^{1/\sigma}} \to 0$ as $x, y \to \infty$ and since $\dfrac{x}{F(x)}$ is bounded above by Propositions 16 and 17, we have

$$\underline{\lim} \frac{\pi(x) \cdot \log x}{F(x)} \geq 1$$

and

$$\overline{\lim} \frac{\pi(x) \cdot \log x}{F(x)} \leq \frac{1}{\sigma} \quad \text{for all } 0 < \sigma < 1.$$

The proposition is therefore proved.

Combining Propositions 16, 17 and 18 we have proved our second theorem:

THEOREM 2. *Let ϕ be an Axiom A flow restricted to a basic set with topological entropy h.*

(i) *If ϕ is topological weak-mixing then*

$$\pi(x) = \{\tau : e^{\lambda(\tau)h} \leq x\} \sim \frac{x}{\log x}.$$

(ii) *If ϕ is not topologically weak-mixing and if a is the least positive eigenfrequency then*

$$\pi(x) \sim \frac{2\pi h/a}{\log x} \sum_{e^{(2\pi n h)/a} \leq x} \cdot \; e^{(2\pi n h)/a}.$$

UNIVERSITY OF WARWICK, COVENTRY, ENGLAND

REFERENCES

[1] L. M. ABRAMOV, On the entropy of a flow, Transl. A. M. S. 49 (1966), 167–170.

[2] V. M. ALEXEEV AND M. V. JACOBSON, Symbolic dynamics and hyperbolic dynamic systems, Phys. Report 75 (1981), 287–325.

[3] R. BOWEN, Equidistribution of closed geodesics, Amer. J. Math. 94 (1972), 413–423.

[4] ———, Periodic orbits for hyperbolic flows, Amer. J. Math. 94 (1972), 1–30.

[5] ———, Symbolic dynamics for hyperbolic flows, Amer. J. Math. 95 (1973), 429–459.

[6] ———, *Equilibrium States and the Ergodic Theory of Anosov Diffeomorphisms*, S.L.N. 470, Springer-Verlag, Berlin, 1975.

[7] D. A. HEJHAL, The Selberg trace formula and the Riemann zeta function, Duke Math. J. 43 (1976), 441–482.

[8] A. MANNING, Axiom A diffeomorphisms have rational zeta functions, Bull. London Math. Soc. 3 (1971), 215–220.

[9] G. A. MARGULIS, Applications of ergodic theory to the investigation of manifolds of negative curvature, Funktional Analiz i Ego Prilozhen, 3 (1969), 89–90.

[10] W. PARRY and S. TUNCEL, *Classification Problems in Ergodic Theory*, L.M.S. Lecture Notes 67, C.U.P., Cambridge, England (1982).

[11] W. PARRY, An analogue of the prime number theorem for shifts of finite type and their suspensions, Israel J. Math. 45 (1983), 41–52.

[12] D. RUELLE, *Thermodynamic Formalism*, Addison-Wesley, Reading, Mass. (1978).

[13] P. SARNAK and A. WOO, Prime geodesic theorems for noncompact quotients (to appear).

[14] JA. G. SINAI, The asymptotic behaviour of the number of closed orbits on a compact manifold of negative curvature, Transl. A. M. S. 73 (1968), 227–250.

[15] N. WIENER, *The Fourier Integral and Certain of its Applications*, C.U.P., Cambridge, England (1967).

(Received February 22, 1983)

ASYMPTOTIC DISTRIBUTION OF CLOSED GEODESICS

Mark Pollicott

§0. INTRODUCTION

In this paper we are concerned with geodesic flows on the unit tangent bundle of compact surfaces of constant negative curvature [2] [12]. One way to study the lengths of closed geodesics is to use the work of Selberg [15], [28]. In this paper we develop a different approach. We apply results for suspension flows with the aid of the symbolic dynamics for geodesic flows due to C. Series [25], [26].

In recent work Parry and the author have derived asymptotic estimates for the number of closed orbits of an Axiom A flow [23]. (In the Axiom A case the proof is heavily dependent on the remarkable work of Bowen [5]). Axiom A flows subsume the case of geodesic flows for surfaces of constant negative curvature. There is a definite loss in generality in studying only geodesic flows. However, in this special case we are able to derive additional results of a more geometric nature. Furthermore, we are able to avoid the complicated machinery of Markov partitions needed for Axiom A flows.

In Section 1 we recall some known results for suspension flows. As an application we relate the unique equilibrium state for a Hölder continuous function (on a subshift of finite type) to the distribution of periodic points. Section 2 contains a short exposition of symbolic dynamics for geodesic flows. In Section 3 we relate the zeta functions for the geodesic flow and the suspension flow associated with it. In Section 4 we recover results of Margulis and Bowen on the distribution of closed geodesics. In the final section we give two new results relating the distribution of closed geodesics to the topology of the surface.

I would like to thank Keith Burns and Caroline Series for helpful comments. I would also like to thank the S.E.R.C. for their financial support.

I am particularly grateful to William Parry for his encouragement and help throughout the course of this work.

§1. CLOSED ORBITS FOR SUSPENDED FLOWS

Let A be an irreducible $k \times k$ zero-one matrix and define

$$\Sigma_A = \{x \in \prod_{-\infty}^{\infty} \{1,\dots,k\} \mid A(x_n, x_{n+1}) = 1 \text{ for all } n \in \mathbb{Z}\}.$$

and $\sigma: \Sigma_A \to \Sigma_A$ by $(\sigma x)_n = x_{n+1}$.

Given $0 < \theta < 1$ we define a metric on Σ_A by $d(x,y) = \theta^n$, where $n \geq 0$ is the largest integer such that $x_i = y_i$, $|i| \leq n-1$.

Assume that $f: \Sigma_A \to \mathbb{R}^+$ is Lipschitz i.e. there exists $C > 0$ s.t. $|f(x) - f(y)| \leq Cd(x,y)$ for all $x, y \in \Sigma_A$. The *suspension space* is defined to be

$$\Sigma_A^f = \{(x,r) \mid x \in \Sigma_A,\ 0 \leq r \leq f(x)\}$$

where $(x, f(x))$ and $(\sigma x, 0)$ are identified.

The *suspension flow* σ^f is given locally by $\sigma_t^f(x,r) = (x, r+t)$, with appropriate identifications. If we assume that $\sup\{h(m) - \int f dm \mid m\ \sigma\text{-invariant}\} = 0$ then the topological entropy $h(\sigma^f)$ is unity. [23] (Here $h(m)$ is the entropy of σ with respect to m.)

Furthermore the measure of maximal entropy for σ^f is $\mu \times \ell / \int f d\mu$, where μ is the unique measure attaining the above supremum and ℓ is linear Lebesgue measure.

The suspended flow is called (topologically) *weak-mixing* if $F \circ \sigma_t^f = e^{iat} F$ has no non-trivial solution ($F \in C(\Sigma_A)$).

A closed σ-orbit $\{x, \sigma x, \ldots, \sigma^{m-1}x\}$ (where $\sigma^m x = x$) corresponds to a closed σ^f-orbit of length $f^m(x) = f(x) + \ldots + f(\sigma^{m-1}x)$.

The following result was proved in [22].

Proposition 1

Let $k : \Sigma_A \to \mathbb{R}^+$ be Lipschitz and σ^f a weak-mixing suspension flow then

$$(1.1) \quad \sum_{m=1}^{\infty} \frac{1}{m} \sum_{\sigma^m x = x} k^m(x) e^{-sf^m(x)} = \frac{\int k d\mu / \int f d\mu}{s-1} + \phi(s)$$

where $\phi(s)$ is analytic on a neighbourhood of $\{s \mid R(s) \geq 1\}$.

This proposition has an immediate interpretation in terms of closed orbit distributions. We write $A(t) \sim B(t)$ if $A(t)/B(t) \to 1$ as $t \to \infty$. We let $O(A(t))$ denote a term whose ratio to $A(t)$ is bounded above.

Proposition 2

Let $k : \Sigma \to \mathbb{R}^+$ be Lipschitz then

$$(i) \quad \sum_{e^{f^m(x)} < t} 1 \sim \frac{t}{\log t}$$

$$(ii) \quad \sum_{e^{f^m(x)} < t} k^m(x) \sim t \int k d\mu / \int f d\mu$$

where σ^f is weak-mixing (summations are over all periodic orbits $\{x, \sigma x, \ldots, \sigma^{m-1}x\}$).

<u>Proof</u>

Denote $F_k(t) = \sum_{e^{f^m(x)}<t} k^m(x)$.

We can rewrite (1.1) as

$$\int_1^\infty t^{-s}\, dF_k(t) = \frac{\int k d\mu / \int f d\mu}{s-1} + \phi(s).$$

The Ikehara-Wiener Tauberian theorem [29] then gives that $F_k(t) \sim t\int k d\mu/\int f d\mu$. This completes the proof of (ii).

For part (i) take $G(t) = \sum_{e^{f^m(x)}<t} 1$. Then

$G(t) = \int_2^t \frac{1}{\log r}\, dF_f(r) + O(1)$. We can rewrite this as follows

$$G(t) = \left[\frac{F_f(r)}{\log r}\right]_2^t + \int_2^t \frac{F_f(r)}{r(\log r)^2}\, dr$$

$$\sim \frac{t}{\log t} + \int_2^t \frac{r}{r(\log r)^2}\, dr$$

$$\sim \frac{t}{\log t} - \left[\frac{r}{\log r}\right]_2^t + \int_2^t \frac{1}{\log t}\, dt$$

$$\sim \quad \ell i(t) \sim \frac{t}{\log t}.$$

(Since $\ell i(t) = \int_2^t \frac{1}{\log t}\, dt \sim \frac{t}{\log t}$).

This completes the proof.

Remarks

(i) By a similar manipulation of Riemann-Stieltjes integrals other asymptotic results can be deduced e.g.

$$\sum_{e^{f^m(x)}<t} e^{-f^m(x)} = \int_2^t \frac{1}{r}\, dG(r) + O(1)$$

$$= \int_2^t \frac{1}{r \log r}\, dr + O(1)$$

$$= \log\log t + O(1).$$

Then following the arguments of [14], §22.18 we can prove

$$\sum_{e^{f^m(x)+f^n(y)}<t} 1 \sim \frac{t}{\log t} \log\log t.$$

(ii) The error terms for the asymptotic estimates depend on the extension of $\phi(s)$ to $R(s) < 1$. For geodesic flows (which we shall see give rise to suspension flows) the error is $O(t^\alpha)$, for some $0 < \alpha < 1$, in Proposition 2 ([15], p. 64). For certain locally constant functions, however, this can never be the case (c.f. [18] Ch. 5 and [24] §5).

(iii) William Parry has shown the author that (1.1) may be generalised to:

$$\sum_{m=1} \frac{1}{m} \sum_{\sigma^m x=x} \frac{(g_1^m \ldots g_k^m)(x)e^{-sf^m(x)}}{[f^m(x)]^{k-1}} = \frac{\int g_1 d\mu \ldots \int g_k d\mu}{[\int f d\mu]^k} \cdot \frac{1}{s-1}$$

$$+ \phi(s)$$

where $g_1, \ldots, g_k \in F_\theta$ and $\phi(s)$ is analytic on a neighbourhood of $\{s | R(s) > 1\}$.

Application (of Proposition 2)

Let μ_0 be the measure of maximal entropy for $\sigma : \Sigma_A \to \Sigma_A$. It is well known that if $Q_0(t) = \{x | \sigma^n x = x, n < t\}$ then $\frac{1}{\text{Card } Q_0(t)} \sum_{x \in Q_0(t)} \delta_x$ converges to μ_0 in the weak* topology. (Here δ_x is the measure consisting of a single atom at x) [1] p. 302.

Let $h : \Sigma_A \to \mathbb{R}$ be Lipschitz and denote its pressure by $P(h)$. By replacing h by $h - P(h)$ we may assume that $P(h) = 0$. Furthermore by adding a coboundary we may assume that $f = -h > 0$. (This requires the Ruelle operator theorem c.f. [23]). If μ is the unique equilibrium state for h then this is unaltered by the above changes.

Assume σ^f is weak-mixing. Define $Q(t) = \{x | \sigma^n x = x, nP(h) - h^n(x) < t\}$ then by Proposition 2, $\frac{1}{\text{Card } Q(t)} \sum_{x \in Q(t)} \delta_x$ converges to μ in the weak* topology.

If σ^f is not weak-mixing then it can be represented (after recoding) by a constant suspension. Therefore the weak* approximation of maximal measures again shows that

$$\frac{1}{\text{Card } Q(t)} \sum_{x \in Q(t)} \delta_x$$ converges to μ.

§2. SYMBOLIC DYNAMICS FOR GEODESIC FLOWS

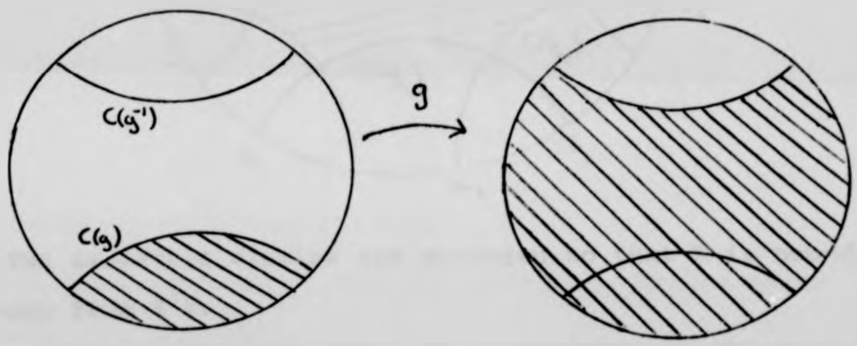
Let M be a compact Riemann surface of curvature $\kappa = -1$ (and genus $g > 2$). Topologically M is a g-holed torus. Define the geodesic flow $\phi_t : T_1M \to T_1M$ on the unit tangent bundle as follows: Given $(x_0, v_0) \in T_1M$ let $\gamma : \mathbb{R} \to M$ be the unique (unit speed) geodesic passing through x_0, in direction v_0, at time $t = 0$ i.e. $\gamma(0) = x_0$ and $\dot{\gamma}(0) = v_0$. Then define $\phi_t(x_0, v_0) = (\gamma(t), \dot{\gamma}(t))$. Thus closed ϕ-orbits in T_1M correspond to closed geodesics in M.

Let D be the Universal covering space for M and $\pi : D \to M$ the projection. Let Γ be the group of covering transformations (i.e. $g \in \Gamma$ is an isometry $g : D \to D$ s.t. $\pi g = \pi$). Since $g > 2$, the Universal cover D is the Poincare disc $\{z \mid |z| < 1\}$ with metric $dx^2 + dy^2 = \dfrac{dr^2}{(1-r^2)^2}$ (where dx, dy are increments in the hyperbolic metric, and dr is an increment in the Euclidean metric) [20] [21].

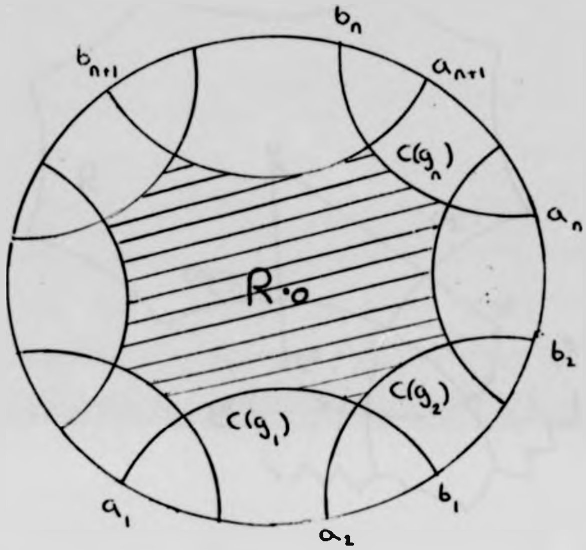
With respect to this metric the geodesics in D are circular arcs meeting S^1 orthogonally [3].

Since Γ is a finitely generated, discontinuous group of isometries on D it forms a Fuchsian group. Elements $g \in \Gamma$ are necessarily linear fractional transformations of the from $g(z) = \dfrac{az + b}{\bar{b}z + \bar{a}}$, where $|a|^2 - |b|^2 = 1$ [13]. Let Γ_0 be a finite set of generators for Γ, together with their inverses.

To find a canonical representation for M in D look at the loci $\{z|\ |g'(z)| = 1\} = C(g)$, where $g \in \Gamma_0$. These are geodesic arcs in D called isometric circles. If $g \in \Gamma_0$ then $gC(g) = C(g^{-1})$ [13]. Furthermore g takes the 'interior' of C(g) to the 'exterior' of $C(g^{-1})$.



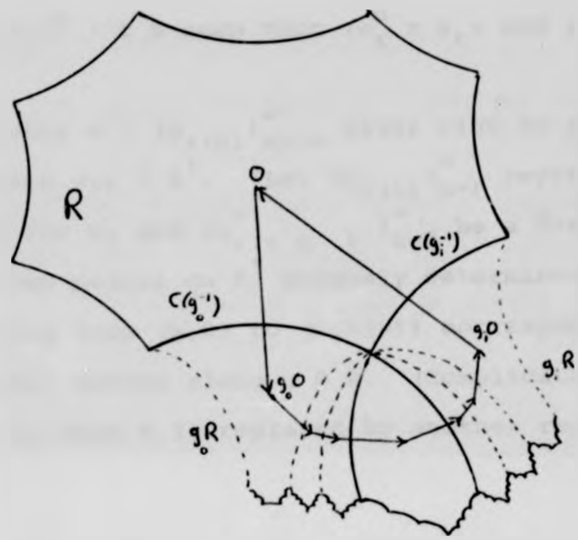
Since Γ is the group of covering transformations we may identify M and D/Γ. The region R exterior to all of the isometric circles C(g), $g \in \Gamma_0$, is called the *Fundamental region*. Since R is a 'maximal' set for which two points are not identified it represents M in the covering space D. (Label the elements of Γ_0 by $g_1, g_2, \ldots$ so that the isometric circles $C(g_1), C(g_2), \ldots$ encircle R in an anti-clockwise direction).



(The isometric circles are arranged so that R is bounded away from S^1).

Other representations of M are given by gR, g $\in$ Γ. Together these 'tile' the disc D by representations all of the same size (in the hyperbolic metric).

By following images of O (under Γ) around any vertex of R we get the unique defining relation for Γ.([13] §27.)



A simplifying assumption is:

(*) For each $g \in \Gamma_0$, $C(g)$ is covered by images of ∂R under Γ. [8], [25].

Consider the piecewise C^2 map $T: S^1 \to S^1$ given by $T(z) = g_i(z)$, $z \in [a_i, a_{i+1})$. (Here $i+1$ is given modulo $4g$). For large N, $|(T^N)'| > 1$ and given $x \in S^1$ we may choose a sequence $(g_{i(n)}^{-1})_{n=0}^{\infty}$ so that $T^n(x) \in [a_{i(n)}, a_{i(n)+1}]$.

Similarly for the map $\bar{T}(z) = g_i(z)$, $z \in (b_i, b_{i+1}]$. Here $y \in S^1$ gives rise to a sequence $(g_{j(n)}^{-1})$ where $\bar{T}^n(y) \in [b_{j(n)}, b_{j(n)+1}]$.

<u>Theorem 1</u> (Series [25])

There exists a subshift (not of finite type) $\Sigma \subseteq \prod_{-\infty}^{+\infty} \Gamma_0$, a Hölder continuous function $h: \Sigma \to \mathbb{R}^+$; and a continuous surjection $\pi:\Sigma^h \to T_1M$ such that $\pi\sigma_t^h = \phi_t\pi$ and π is bounded-to-one.

A sequence $w = (g_{r(n)})_{n=-\infty}^{\infty}$ gives rise to expansions for two points $x,y \in S^1$. (Let $(g_{r(n)})_{n=1}^{\infty}$ represent a T-expansion for x, and $(g_{r(-n)}^{-1})_{n=0}^{\infty}$ be a $\bar{T}$-expansion for y). These two points on S^1 uniquely determine a geodesic γ on D. Flowing from $(w,0)$ to $(w,h(w))$ corresponds to the unit tangent vector moving along $\gamma \cap \bar{R}$. (Complications arise if $\gamma \cap \bar{R} = \phi$, when R is replaced by another representation gR, $g \in \Gamma$).

Note

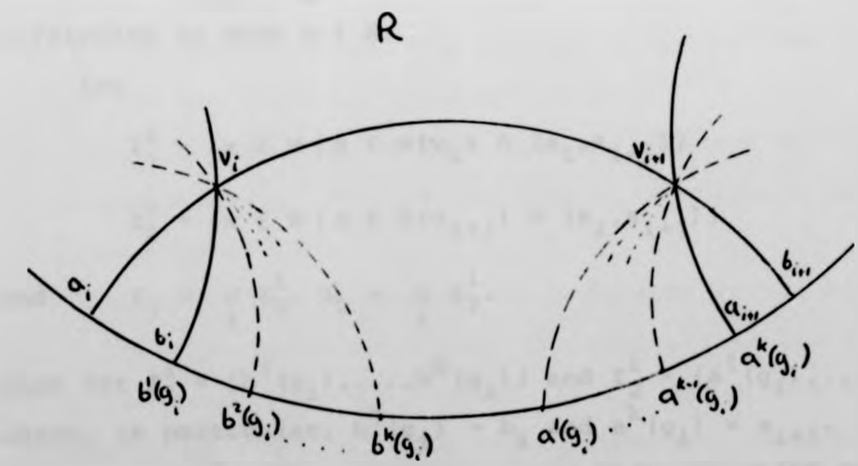
Series shows that the condition (*) is not essential and a flow which satisfies this condition can be related to one which does not by a quasi-conformal mapping ([25] §6).

The measure of maximal entropy for ϕ is the Riemann measure ([6] p. 420). Furthermore the geodesic flow ϕ is topologically weak-mixing [11].

§3. CLOSED ORBITS AND ZETA FUNCTIONS

To use the results of Section 1 we need to replace Σ by a subshift of finite type Σ_A.

Bowen and Series constructed a Markov partition for $T:S^1 \to S^1$ as follows: For each vertex v_i of R let $N(v_i)$ be those geodesic arcs passing through v_i which are images of $C(g)$ $(g \in \Gamma_0)$ under Γ [8]



Let W be the points where arcs in $\bigcup_i N(v_i)$ meet S^1. If we still assume (*) then $\{a_i\}, \{b_i\} \subseteq W$. Furthermore $TW \subseteq W$ (here we could take $T(a_i)$ to be either $g_i(a_i)$ or $g_{i-1}(a_i)$ and this would still be true) [8].

The points W therefore divide S^1 up into a Markov partition. Let Σ_A^+ be the one-sided shift of finite type given by this Markov partition. Since this new partition refines the old one there is an obvious map $i:\Sigma_A^+ \to \Sigma^+$ (where Σ^+ is the one-sided subshift corresponding to Σ). This extends to $i:\Sigma_A \to \Sigma$. We want to compare closed orbits in $\Sigma_A^{h\circ i}$ and T_1M. There will be a one-one correspondence unless a T_1M closed orbit has an Σ_A expansion which is not unique. By periodicity it suffices to compare closed orbits for $\sigma:\Sigma_A^+ \to \Sigma_A^+$ and $T:S^1 \to S^1$. In this case the only difficulty is when $x \in W$.

Let

$$I_1^i = \{x \in W \mid x \in N(v_i) \cap (a_i,a_{i+1}]\}$$

$$I_2^i = \{x \in W \mid x \in N(v_{i+1}) \cap [a_i,a_{i+1}]\}$$

and $$I_1 = \bigcup_i I_1^i, \; I_2 = \bigcup_i I_2^i.$$

Next let $I_1^i = \{b^1(g_i),\ldots,b^k(g_i)\}$ and $I_2^i = \{a^1(g_i),\ldots,a^k(g_i)\}$, where, in particular, $b^1(g_i) = b_i$ and $a^k(g_i) = a_{i+1}$. The elements of I_1^i, I_2^i are listed in the order in which they are labelled around S^1. (Here $k = 2g-1$). Because of these definitions $g_i\, I_1^i \subseteq I_1$ and $g_i\, I_2^i \subseteq I_2$. Furthermore

$$\begin{cases} Ta^r(g_i^{-1}) = a^{r-1}(g_j) & \text{(for some } j = j(r,i)\text{)} \\ Tb^r(g_i^{-1}) = b^{r+1}(g_{j'}) & \text{(for some } j' = j'(r,i)\text{)} \end{cases}$$

(In fact, $j(r,i) = i - 1$ ($r \neq 1$), $j(1,i) = i - 2$ and $j'(r,i) = i + 1$ ($r \neq k$), $j'(k,i) = i + 2$).

where j , j' and r,r-1 and r+1 are given modulo $4g$ and modulo k respectively

Note

The canonical case is where the isometric circles have the clockwise ordering

$$g_{4g} \cdot \cdot \cdot g_4 g_3 g_2 g_1$$
$$= p_1 q_1 p_1^{-1} q_1^{-1} \ldots p_g q_g p_g^{-1} q_g^{-1} \quad \text{(c.f. [21])}.$$

In this case g_{i+1} follows g_i^{-1} in the cyclic sequence

$$q_g^{-1} \; p_g q_g p_g^{-1} \ldots q_1^{-1} p_1 q_1 p_1^{-1} .$$

(These two sequences are closely connected with the surface symbol and the defining relation [21]).

If g is even then we have that I_1 contains $2g$ T-orbits of period $2(2g-1)$. If g is odd then I_1 contains 4 orbits of period $g(2g-1)$.

Points in I_1 lie between two intervals in the Markov partition. A consistent choice of intervals to the left or to the right for each T-orbit (in I_1) gives rise to twice the number of σ-orbits. The set I_2, however, does not give rise to any additional orbits in Σ_A. This is because the choice of intervals to the right breaks down at a_i. Thus the primitive closed orbits for ϕ and $\sigma^{h \circ i}$ differ by only a finite number ℓ of orbits of length C (corresponding to σ-orbits of period p), say. (Here p and ℓ depend only on g, the genus of the surface. The actual value of these constants will prove to be unimportant.).

Comparing zeta functions we see that

$$\zeta_{\psi}(s) = \prod_{\tau} (1 - e^{-s\lambda(\tau)})^{-1}$$

(Here the product is over closed orbits τ of length $\lambda(\tau)$)

$$= \exp \left(\sum_{m=1}^{\infty} \frac{1}{m} \sum_{\sigma^m x = x} e^{-sh\circ i^m(x)} - \ell \sum_{n=1}^{\infty} \frac{e^{-snC}}{n} \right)$$

$$= \zeta_{\sigma^{h\circ i}}(s) \exp - \ell \log (1-e^{-sC})$$

$$= \zeta_{\sigma^{h\circ i}}(s)(1 - e^{-sC})^{\ell}$$

Remarks

(i) Let $G(z)$ be the generating function for homotopy class word lengths then:

$$G(z) = \sum_{n=1}^{\infty} z^n . \text{ Card } \{\text{Homotopy classes length } n\}.$$

Cannon has shown that $G(z)$ is rational [10]. Let us consider the generating function for free homotopy classes $G_0(z)$. Every free homotopy class corresponds to exactly one closed geodesic ([4], VII.6). Thus

$$G_0(z) = \sum_{n=1}^{\infty} z^n \text{ Card } \{\text{Free homotopy classes length } n\}$$

$$= \sum_{n=1}^{\infty} \frac{z^n}{n} \text{ Card } \{x \mid \sigma^n x = x\} - \ell \sum_{n=1}^{\infty} \frac{z^{np}}{n}$$

$$= \log \left[\frac{(1-z^p)^{\ell}}{\det(I-zA)} \right] .$$

In particular, $\exp G_\sigma(z)$ is rational

(ii) Let $\alpha > 0$ be the growth rate of cyclically reduced words i.e. those corresponding to closed geodesics. Then e^α is a pole for $G_\sigma(z)$ and so a zero of the polynomial $\det(I-zA)$. Thus e^α is an algebraic number.

Notice that since ϕ and $\sigma^{h \circ i}$ have closed orbits of the same lengths $\sigma^{h \circ i}$ is weak-mixing [7].

§4. ASYMPTOTIC RESULTS FOR CLOSED GEODESICS

In [16], Hejhal gives an asymptotic formula for the number of closed geodesics on a surface of constant negative curvature. We shall now give a proof with a more dynamic flavour.

Let $\nu(t)$ be the number of closed geodesics of length at most t. Then $\nu(t) = \nu_\phi(t)$, the number of closed ϕ-orbits of period at most t.

<u>Theorem 2</u> (Margulis, c.f. Hejhal [15], [16])

Let M be a compact surface of curvature $\kappa = -1$ then $\nu(t) \sim \frac{e^t}{t}$.

<u>Proof</u>

We saw in the previous section that $\zeta_\phi(s) = \zeta_{\sigma^{h \circ i}}(s) . (1-e^{-sC})^\ell$ and in particular there are only ℓ additional primitive closed orbits for $\sigma^{h \circ i}$ compared with ϕ (each of length C). Thus $\nu_\phi(t)$ and $\nu_{\sigma^{h \circ i}}(t)$ differ only by iterates of these orbits i.e. $\nu_{\sigma^{h \circ i}}(t) = \nu_\phi(t) + \ell[t/C]$.

By Proposition 2 $\nu_{\sigma^{h \circ i}}(t) \sim e^t/t$ thus $\nu_\phi(t) \sim e^t/t$. This completes the proof.

We can also approach the problem of distribution of closed orbits on T_1M. Let B be a subset of T_1M whose boundary has measure zero. Let $\nu_B(t)$ be the total sojourn time in B of closed orbits whose periods are at most t.

<u>Theorem 3</u> (Bowen [6], cf. Parry [22])

Let M be a compact surface of curvature $\kappa = -1$, and let $B \subseteq T_1M$ have $m(\partial B) = 0$, then $v_B(t) \sim \frac{m(B)}{m(M)} e^t$.

<u>Proof</u>

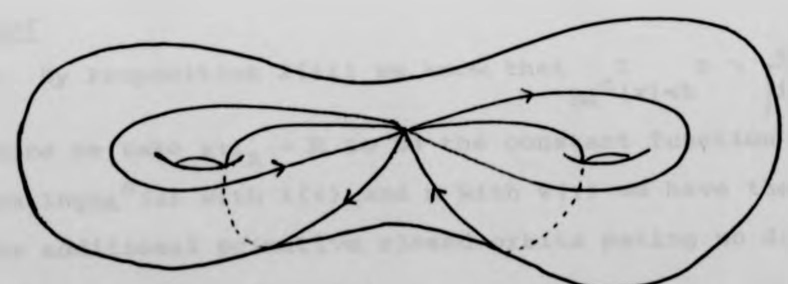
We already know this result is true for suspension flows and that the scalar $m(B)/m(M)$ is appropriate. Again the only difference is due to a finite number of orbits. This contributes a difference of order t. Therefore the result follows.

<u>Remark</u>

The above theorem shows that closed orbits are uniformly distributed in T_1M (with respect to m). This means that closed geodesics are uniformly distributed on M (with respect to Riemann measure) and in direction (with respect to Lebesgue measure).

§5. FURTHER DISTRIBUTION RESULTS

By following the formulation due to C. Series we are able better to describe the distribution of closed orbits with respect to the topological structure. Under the assumption (*) the sides of R form closed loops from arcs of closed geodesics. Topologically these loops ring the holes in the surface or pass through such holes. We can interpret a geodesic as passing through a hole if it cuts the first type of loop. Similarly we may interpret a geodesic as circling a hole if it traverses the second type of loop.



For a closed geodesic τ let $w(\tau)$ be the total number of such crossings by τ . Assume that the corresponding σ^{hw}-orbit lies over a σ-periodic point $\sigma^n x = x \in \Sigma_A$. Then the period of x is the number of crossing of R, say, i.e. $w(\tau) = n$. For $g \in \Gamma_0$ let $w_g(\tau)$ be the number of crossings of the loop corresponding to $C(g) \cap \bar{R}$.

<u>Theorem 4</u>

Let M be a compact surface with curvature $\kappa = -1$ and satisfying (*) then

(i) There exists $C > 0$ such that $\sum_{\lambda(\tau)<t} w(\tau) \sim Ce^t$.

(ii) For $g \in \Gamma_0$ there exists $C_g > 0$ such that

$$\sum_{\lambda(\tau)<t} w_g(\tau) \sim C_g . e^t .$$

<u>Proof</u>

(i) By Proposition 2(ii) we know that $\sum_{h^n(x)<t} n \sim \frac{e^t}{\int f d\mu}$ (where we take $k : \Sigma_A \to \mathbb{R}$ to be the constant function 1). Equating $h^n(x)$ with $\lambda(\tau)$ and n with $w(\tau)$ we have the result. (The additional primitive closed orbits making no difference to the asymptotic formula).

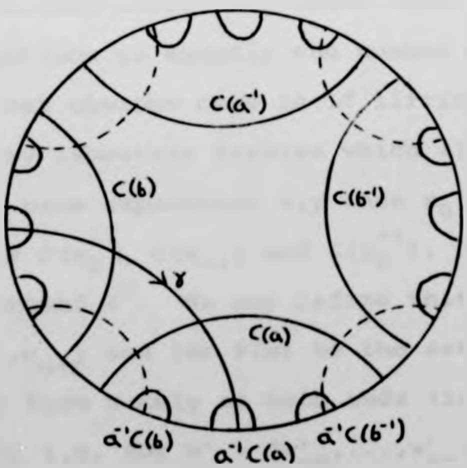
(ii) Let $B_g \subseteq \Sigma$ be those sequences representing geodesics crossing $C(g) \cap \bar{R}$, say. Let $C_g = \mu(B_g)/\int h d\mu$. The result follows from Proposition 2(ii) by taking $k = \chi_{B_g}$ (after extending Proposition 2 to $k = \chi_{B_g}$ by approximation.).

Remark

If $C(g) \cap \bar{R}$ has length $\ell(g)$ then $C(g) = \ell(g)/\pi.\mathrm{vol}M$. If ∂R has length $\ell(\partial R)$ then $C = \ell(\partial R)/\pi.\mathrm{vol}M$.

The above theorem relates lengths of closed geodesics to their intersections with fixed geodesics $C(g)$. We want to adapt these ideas to accommodate self-intersections.

Let M be a manifold of infinite volume where generators in Γ_0 satisfy $C(g_i) \cap C(g_j) = \phi$ with $g_i, g_j \in \Gamma_0$, $(g_i \neq g_j)$. [17], [27].



A geodesic γ in D is labelled by a bi-infinite sequence from Γ_0. Let γ cut $C(g_0)$, $g_0^{-1}C(g_1)$, $g_0^{-1}g_1^{-1}C(g_2)$, $g_0^{-1}g_1^{-1}g_2^{-1}C(g_3)$,... where we start in R. Furthermore, going in the opposite direction assume γ cuts $C(g_{-1}^{-1})$, $g_{-1}\, C(g_{-2}^{-1})$, $g_{-2}g_{-1}\, C(g_{-3}^{-1})$,... We represent γ by the sequence (g_n^{-1}) [27].

If the geodesic γ lies in the non-wandering set Ω (which includes all closed geodesics) then the sequence does not terminate. The only admissibility condition is that g is not followed by g^{-1}. Admissible sequences form a shift of finite-type Σ_A. For a geodesic γ with a sequence $x \in \Sigma_A$ let $r(x)$ be the hyperbolic length of $\gamma \cap \bar{R}$. Then $r:\Sigma_A^+ \to \mathbb{R}^+$ is Hölder continuous. There is a topological conjugacy $\pi:\Sigma_A^r \to T_1\Omega$ s.t. $\pi\sigma_t^r = \phi\pi$.

In order that a geodesic on M intersects itself it must have distinct lifts to D which cross. The number of self-intersections is exactly the number of such intersections in D. The most obvious case is if liftings γ and γ' enter and leave R by isometric circles which alternate around S^1 i.e. if γ,γ' have expansions x,y then $x_0 \neq y_0$ and $x_{-1} \neq y_{-1}$ and the pairs $C(x_0^{-1})$, $C(x_{-1})$ and $C(y_0^{-1})$, $C(y_{-1})$ separate each other around S^1. We may refine this: Let $w = [w_{-n},\ldots,w_{n-1}]$ and let P(w) be the set of 2n-cylinders which differ from w only at both ends (in -n and n-1 co-ordinates) i.e. for $w' = [w'_{-n},\ldots,w'_{n-1}] \in P(w)$; $w'_{-n} \neq w_{-n}$, $w'_{n-1} \neq w_{n-1}$, $w_i = w'_i$ $(-n < i < n-1)$. Furthermore, impose the condition on P(w) that the circular arcs $w_0 w_1 \ldots C(w_{n-1}^{-1})$, $w_{-1}^{-1} w_{-2}^{-1} \ldots C(w_{-n})$ separate $w'_0 w'_1 \ldots C(w'^{-1}_{n-1})$ and $w'^{-1}_1 \, w'^{-1}_2 \ldots C(w'_{-n})$ (c.f. [8]).

If x defines a geodesic then a self-intersection occurs if $x \in w$ and $\sigma^k x \in w' \in P(w)$. We can count the number of such self-intersections (given $\sigma^m x = x$) as $S_n(x) =$ $\frac{1}{2}\sum_{w} \sum_{w' \in P(w)} [\chi_w]^m(x) . [\chi_{w'}]^m(x)$, where χ_w is a characteristic function.

The number of self-intersections per unit length is $S_n(x)/r^m(x)$. Consider

$$\eta(s) = \sum_{\substack{w \\ w' \in P(w)}} \left\{ \sum_{m=1}^{\infty} \frac{1}{m} \sum_{\sigma^m x = x} \frac{[\chi_w]^m(x)\,[\chi_{w'}]^m(x)}{2 . r^m(x)} e^{-sr^m(x)} \right\}$$

$$= \frac{\ell_n}{s-1} + \psi(s) \text{ where } \ell_n = \frac{1}{2\int r d\mu} \sum_{\substack{w \\ w' \in P(w)}} \mu(w).\mu(w').$$

Thus an nth approximation to counting self-intersections of closed orbits is $\sum_{\lambda(\tau)<t} \frac{S_n(\tau)}{\lambda(\tau)} \sim \ell_n e^t$.

We now consider the limit of these approximations. Let $S(\tau)$ be the number of self-intersections of τ. Define $B \subseteq (S^1 \times S^1) \times (S^1 \times S^1)$ by

$B = \{(x_1, x_2; y_1, y_2) : x_i < y_j < x_k < y_\ell$ lie in order around S^1;

$\{i,k\} = \{j,\ell\} = \{1,2\}$; and the corresponding geodesics intersect in R $\}$

Theorem 5

$$\sum_{\lambda(\tau)<t} \frac{S(\tau)}{\lambda(\tau)} \sim (\pi\ \mu \times \pi\ \mu)(B).e^t/2$$

This follows by approximation. Thus the average number of self-intersections per unit length is $(\pi\ \mu \times \pi\ \mu)(B)/2$.

REFERENCES

[1] V.M. Alexeev and M.V. Jacobson, Symbolic dynamics and hyperbolic dynamic systems, Phys. Report 75 No. 5 (1981) 287-325.

[2] V.I. Arnold and A. Avez, *Ergodic problems of Classical Mechanics*, Benjamin, New York (1968).

[3] A.F. Beardon, "Geometry of Discrete Groups" in *Discrete groups and Automorphic functions* (ed. W.J. Harvey), Academic, London, (1977).

[4] M. Berger, *Lectures on Geodesics in Riemannian Geometry*, Tata Institute, Bombay 1965.

[5] R. Bowen, Symbolic dynamics for hyperbolic flows, Amer. J. Math. 95 (1973) 429-459.

[6] R. Bowen, Equidistribution of closed goedesics, Amer. J. Math. 94 (1972) 413-423.

[7] R. Bowen, Mixing Anosov flows, Topology 15 (1976) 77-79.

[8] R. Bowen and C. Series, Markov maps associated with Fuchsian groups, Publ. Math. I.H.E.S. 50 (1979) 153-170.

[9] J.S. Birman and C. Series, An algorithm for simple curves on surfaces (Preprint).

[10] J.W. Cannon, The growth of the closed surface groups and the compact hyperbolic Coxeter groups. (Preprint).

[11] P. Eberlein, Geodesic flows on negatively curved manifolds II, Trans. Amer. Math. Soc, 178 (1973) 57-82.

[12] S.V. Fomin and I.M. Gelfand, Geodesic flows on manifolds of constant negative curvature, Transl. Amer. Math. Soc. (2), 1 (1955) 49-65.

[13] L.R. Ford, *Automorphic functions*, McGraw Hill, New York, (1929).

[14] G.H. Hardy and E.M. Wright, An introduction to the theory of numbers, Oxford University Press, Oxford (1938).

[15] D.A. Hejhal, The Selberg trace formula for PSL $(2,\mathbb{R})$ Vol. I, SLN 548, Springer, Berlin (1976).

[16] D.A. Hejhal, The Selberg trace formula and the Riemann zeta function, Duke Math. J., 43 (1976), 441-482.

[17] E. Hopf, Fuchsian groups and ergodic theory, Trans. Amer. Math. Soc., 39 (1936) 299-314.

[18] A.E. Ingham, The distribution of prime numbers, Cambridge Mathematical Tracts No. 30, Cambridge, (1932).

[19] A.B. Katok, Ya. G. Sinai and A.M. Stepin, Theory of dynamical systems and general transformation groups with invariant measure, J. of Soviet Math. 7 (1977) 974-1065.

[20] J. Lehner, *A short course in automorphic functions*, Holt, Rinehart and Winston, New York, (1965).

[21] A.M. Macbeath, Discontinuous groups and birational transformations (Lecture Notes).

[22] W. Parry, Bowen's equidistribution theory and the Dirichlet Density Theorem (To appear in Ergod. Th. and Dynam. Sysm.)

[23] W. Parry and M. Pollicott, An analogue of the prime number theorem for closed orbits of Axiom A flows (To appear in Annals of Math.).

[24] M. Pollicott, A complex Ruelle-Perron-Frobenius theorem and two counter examples (To appear in Ergod. Th. and Dynam. Sys.)

[25] C. Series, Symbolic Dynamics for geodesic flows, Acta Math., 146 (1981) 103-128.

[26] C. Series, On coding geodesics with continued fractions, *Ergodic Theory* (Sem. Plans-sur-Box), Monograph Enseign Math., 29, Univ. Geneve, Geneva (1981).

[27] C. Series, The infinite word problem and limit sets in Fuchsian groups, Ergod. Th. and Dynam. Sys., 1 (1981) 337-360.

[28] A.B. Venkov, Spectral theory of automorphic functions, the Selberg zeta-function, and some problems of analytic number theory and mathematical physics, Russian Math. Surveys 34 (3) (1979) 79-153.

[29] N. Wiener, The Fourier integral and certain of its applications.

APPENDIX: A SURVEY OF SYMBOLIC DYNAMICS FOR AXIOM A FLOWS

§0. REFLECTIONS ON AXIOM A FLOWS

This appendix is included for the benefit of the reader and is purely expository. In a number of places we have omitted rigorous proofs in the interests of brevity. References to the original material are given where appropriate.

Axiom A flows are generalisations of Anosov flows that were first introduced by Smale [13] (Anosov flows include the special case of geodesic flows on surfaces of constant negative curvature ([1] §14)). For the Axiom A case interest is centred on the basic sets. These contain the most interesting activity and in particular contain the periodic orbits. However they need not have any of the differential structure of the manifold in which they are contained. We shall try to emphasize the different properties of the manifold and the basic set when used.

The study of Axiom A flows was greatly advanced by the work of Bowen [2], [3], [4]. Bowen showed that such flows could be very closely modelled by suspension flows. In terms of the maximal measures this correspondence is a measure isomorphism conjugating the flows.

Topologically there is 'almost' a conjugacy of the flows (the difficulty being confined to a 'thin' set). Furthermore a combinatorial result (Manning's lemma ([2] §5)) actually permits numbers of closed orbits to be compared for each flow.

The device used to actually construct the suspension flow is a Markov 'partition' on transverse sections to the flow ([2], p.436). By incorporating a number of Bowen's most elegant arguments from the diffeomorphism case we show that their existence is a natural consequence of two important properties of the flow - expansiveness and the tracing property. (cf. [5] §3.C).

After constructing the symbolic flow the asymptotic growths of the numbers of closed orbits are compared.

We finally conclude with an outline of Manning's lemma applied in this context ([2] §5).

§1. AXIOM A FLOWS

1.1 Definitions

Let $\phi_t : M \to M$ be a C^1 flow on a compact manifold M. A compact invariant set Λ, containing no fixed points, is called *hyperbolic* if the tangent bundle restricted to Λ can be written as the Whitney sum of three $D\phi_t$-invariant continuous sub-bundles

$$T_\Lambda M = E \oplus E^s \oplus E^u$$

where E is the one-dimensional bundle tangent to the flow, and there are constants $C, \lambda > 0$ so that:

(a) $\|D\phi_t(v)\| \leq Ce^{-\lambda t}\|v\|, \quad v \in E^s, \ t \geq 0$

(b) $\|D\phi_{-t}(v)\| \leq Ce^{-\lambda t}\|v\|, \quad v \in E^u, \ t \geq 0.$

The *non-wandering set* Ω is defined by

$$\Omega = \{x \in M : \text{for every neighbourhood } V \text{ of } x,\ t_0 > 0,\ \exists t > t_0 \text{ s.t. } \phi_t V \cap V \neq \emptyset\}.$$

A closed ϕ_t-invariant set $\Lambda \subseteq \Omega$ is called *basic* if

(a) Λ contains no fixed points and is hyperbolic

(b) The periodic orbits in Λ are dense in Λ

(c) $\phi_t|_\Lambda$ is topologically transitive

(d) $\Lambda = \bigcap_{t \in \mathbb{R}} \phi_t U$, for some open set $U \supseteq \Lambda$.

The flow ϕ is said to satisfy *Axiom* A if Ω is a finite disjoint union of basic sets (together with a finite number of isolated fixed points.) ([13] p. 803).

1.2 Canonical Co-ordinates

Given $x \in \Lambda$ and $\varepsilon > 0$ define:

$$W^s_\varepsilon(x) = \{y \in M : d(\phi_t x, \phi_t y) < \varepsilon, \ (t > 0);$$

$$d(\phi_t x, \phi_t y) \to 0 \text{ as } t \to \infty\}$$

$$W^u_\varepsilon(x) = \{y \in M : d(\phi_{-t} x, \phi_{-t} y) < \varepsilon, \ (t > 0);$$

$$d(\phi_{-t} x, \phi_{-t} y) \to 0 \text{ as } t \to \infty\}$$

It transpires that these are C^1 submanifolds of M satisfying: (for small $\varepsilon > 0$)

$$\phi_t \, W^s_\varepsilon(x) \subseteq W^s_{\varepsilon C e^{-\lambda t}}(\phi_t x) \qquad (t > 0)$$

$$\phi_{-t} \, W^u_\varepsilon(x) \subseteq W^u_{\varepsilon C e^{-\lambda t}}(\phi_{-t} x) \qquad (t > 0)$$

[10], [11].

The significance of the stable and unstable manifolds is that they induce *canonical co-ordinates* i.e. for small $\varepsilon > 0$, there exists $\delta > 0$ such that if $x, y \in \Lambda$ and $d(x,y) < \delta$ then there exists a unique $v = v(x,y) \in \mathbb{R}$ with $|v| < \varepsilon$ and

$$W^s_\varepsilon(\phi_v x) \cap W^u_\varepsilon(y) \neq \emptyset.$$

Furthermore, this intersection is a single point $\langle x,y\rangle \in \Lambda$ [13] [12].

[The maps v, $\langle\ ,\ \rangle$ are continuous on

$$\{(x,y) \in \Lambda \times \Lambda : d(x,y) < \delta\}].$$

§2. TWO IMPORTANT PROPERTIES OF AXIOM A FLOWS

2.1 Expansiveness

We call a continuous flow ψ on a compact metric space X *expansive* if $\forall\varepsilon > 0$, $\exists\delta > 0$, such that whenever $d(\psi_t x, \psi_{s(t)} y) < \delta$ (all $t \in \mathbb{R}$, some $x,y \in X$ and some continuous, monotonically increasing function $s:\mathbb{R} \to \mathbb{R}$ with $s(0) = 0$) then $y = \phi_v x$ for some $|v| < \varepsilon$.

Remark

This can be interpreted loosely as saying that: "If the orbits of x and y are sufficiently close, for all time, then x and y share the same orbit".

Proposition (Bowen ([4] p.4))

For an Axiom A flow ϕ (restricted to a basic set Λ) $\phi_t: \Lambda \to \Lambda$ is flow expansive

2.2 The tracing property

Let ψ be a flow on a compact metric space X.

For ε, $T > 0$ an (ε,T)-*chain* is a pair of bi-infinite sequences $((x_n), (t_n))$ satisfying $t_n > T$ and $d(\phi_{t_n} x_n, x_{n+1}) < \varepsilon$ (all $n \in \mathbb{Z}$).

We say a point $x \in X$ δ-*traces* an (ε,T)-chain if there exists a continuous monotonically increasing function $s:\mathbb{R} \to \mathbb{R}$ with $s(0) = 0$ such that:

(i) $d(\phi_{s(t)}x, \phi_{t-\alpha(n)}x_n) <$

where $\alpha(n) = \sum_{i=0}^{n-1} t_i < t < \alpha(n+1)$ and

(ii) $d(\phi_{s(t)}x, \phi_{t+\beta(n)}x_n) <$

where $\beta(n) = \sum_{i=-n}^{-1} t_i > t \geqslant \beta(n-1)$.

<u>Proposition</u> (Franke and Selgrade [9])

For an Axiom A flow ϕ (restricted to a basic set Λ) and T, $\delta > 0$ there is an $\varepsilon > 0$ such that every (ε,T)-chain in Λ can be δ-traced.

<u>Remark</u>

Intuitively this means that fragments of orbits can be approximated by a 'complete' orbit.

<u>Remark</u>

Bowen and Walters have shown that expansiveness is sufficient to model the flow by a suspension (not necessarily Hölder continuous) over a subshift (not necessarily of finite type) [8].

<u>Remark</u>

For Axiom A diffeomorphisms Bowen showed that the Markov partition arises in a very elegant way as a consequence of an appropriate tracing property [5].

§3. DIFFERENTIAL TOPOLOGY (ON THE MANIFOLD)

In this section we confine ourselves to the structure of the actual manifold M. Proofs of results in this section can be found in Section 7 of [2].

Given $\alpha > 0$, choose $D_1, \ldots, D_n$ differentiable closed discs transverse to the flow satisfying

(i) diam $D_i < \alpha$

(ii) for $i \neq j$, either $D_i \cap \phi_{[0,4\alpha]} D_j = \phi$ or

$$D_j \cap \phi_{[0,4\alpha]} D_i = \phi.$$

(c.f. [2] p. 455).

The lemmas in this section are consequences of the C^1 structure of ϕ.

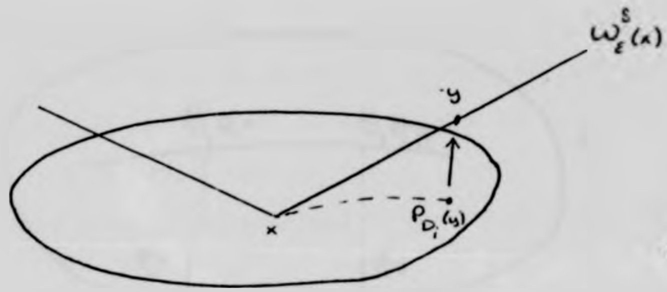
Since D_i, $W^s_\varepsilon(x)$, $W^u_\varepsilon(x)$ are all C^1 we have the following:

Lemma 1 (Bowen)

There exists $C_1 > 0$ such that for $x \in D_i \cap \Lambda$ and $y \in W^s_\varepsilon(x) \cap \phi_{[-\alpha,\alpha]} D_i$ (or $y \in W^u_\varepsilon(x) \cap \phi_{[-\alpha,\alpha]} D_i$) then

$$d(P_{D_i}(y), x) \geq C_1 d(y,x)$$

(where $P_{D_i}(y)$ is the projection of y onto D_i).



The above lemma is used to prove the following:

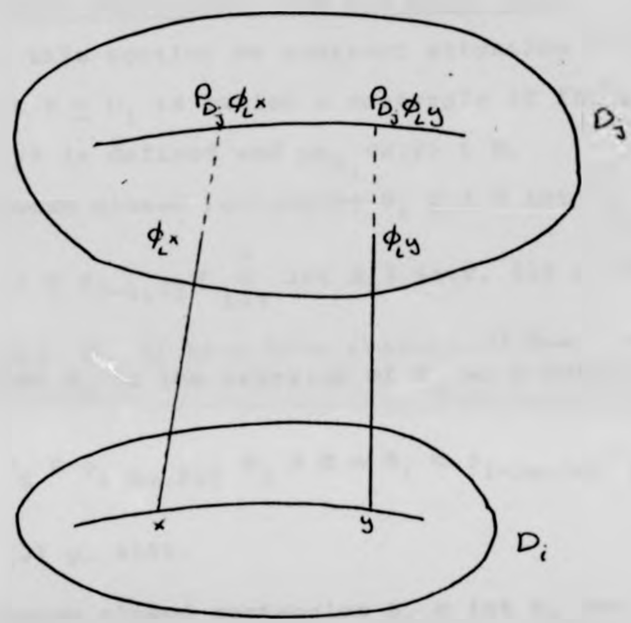
Lemma 2 (Bowen)

For $0 < \tau < 1$ there exists L such that for $x \in D_i$, $y \in W^s(x,D_i)$ with $\phi_L(x), \phi_L(y) \in \phi_{[-2\alpha,2\alpha]}D_j$ then

$$d(P_{D_j}\phi_L(x), P_{D_j}\phi_L(y)) < \tau d(x,y)$$

(and similarly for ϕ_{-L}) where $W^s(x,D_i) = P_{D_i}W_\varepsilon^s(x)$.

Because $\phi:M \to M$ is C^1 then the flow satisfies certain Lipschitz conditions



<u>Lemma 3</u>

Given $L > 0$, there exists $C_2 > 0$ such that $x,y \in D_i$ and $|t| < 2L$ means $d(\phi_t x, \phi_t y) < C_2 d(x,y)$.

<u>Lemma 4</u> (Bowen)

There exists C_3 such that if $x,y \in D_i$ and $\phi_t x \in D_j$, $\phi_t y \in \phi_{[-2\alpha,2\alpha]} D_j$ then $d(\phi_t x, P_{D_j}\phi_t y) < C_3 d(x,y)$.

§4. MARKOV PARTITIONS (FOR THE BASIC SET)

In this section we restrict attention to $\phi_t: \Lambda \to \Lambda$.
A subset $R \subseteq D_i$ is called a *rectangle* if for all $x,y \in R$, $pr_{D_i}\langle x,y\rangle$ is defined and $pr_{D_i}\langle x,y\rangle \in R$.

Choose closed rectangles $B_i \subseteq \Lambda \cap \text{int } D_i$ s.t.

(iii) $\Lambda \subseteq \phi_{[-\alpha,0]}\left(\bigcup_{i=1}^{n} \text{int } B_i\right)$ (c.f. [2] p. 455).

The D_i's should have been chosen with this in mind.

(Here int B_i is the interior of B_i as a subset of $\Lambda \cap \text{int } D_i$).

(iv) $B_i \cap \phi_{[-2\alpha,2\alpha]} B_j \neq \emptyset \Rightarrow B_i \subset \phi_{[-3\alpha,3\alpha]} D_j$

(c.f. [2] p. 455).

Choose closed rectangles $K_i \subseteq \text{int } B_i$ and $\delta > 0$ s.t.

(v) Any subset of Λ of diameter at most 3δ lies in some $\phi_{[-2\alpha,2\alpha]}K_i$. (c.f. [2] p. 456).

4.1 Maps between sections

We would have liked the K_i to be used as a starting point in constructing the Markov 'partition'. Unfortunately, the discontinuity of the Poincaré map makes this difficult. The first step, therefore, is to define a family of maps which are only locally defined, but are continuous on their domains. (c.f. [2] p. 456).

Cover each K_i by a finite family $\mathcal{V}_i$ of closed sets s.t. diam $\phi_t V < \delta$ for all $|t| < 2L$, $V \in \mathcal{V}_i$.

By (v):

$$B_\delta(\phi_{-L}V) \subseteq \phi_{[-2\alpha,2\alpha]}K_{a(v)}$$

$$B_\delta(\phi_L V) \subseteq \phi_{[-2\alpha,2\alpha]}\ K_{b(v)}$$

(for some $a(V)$, $b(V)$).

Let ε_0 be the Lebesgue number of the covers V_i and take $\varepsilon_1 << \varepsilon_0$ (to be made precise later).

Cover the K_i's by ε_1-balls (in int B_i) i.e. $\{B_{\varepsilon_1}(x_n)\}_{n=1}^N$ where $(x_n)_{n=1}^N \subseteq B_i$.

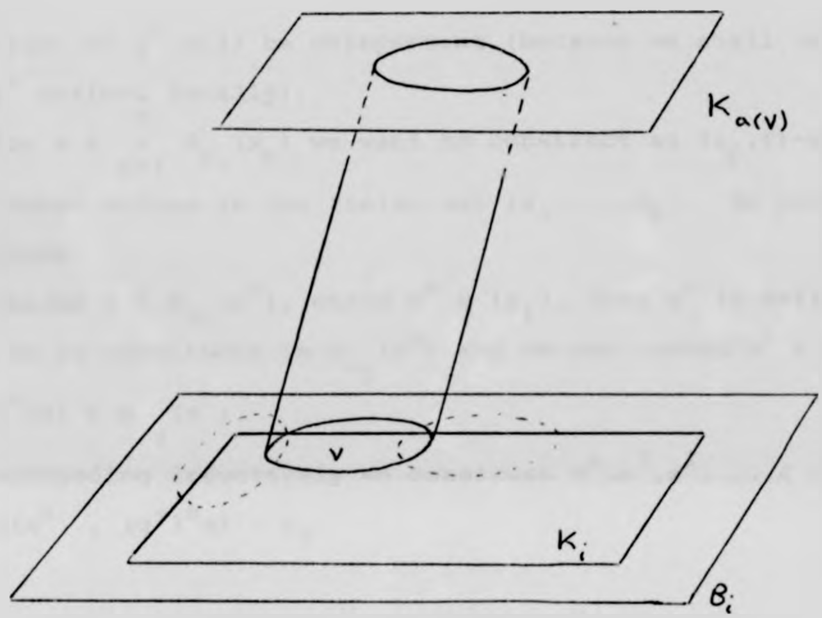
Because of this set up: Given x_n choose $V \in \bigcup_i \mathcal{V}_i$ such that $B_{\varepsilon_1}(x_n) \subseteq V$ and define

$$g_n^+ : B_{\varepsilon_1}(x_n) \to K_{a(v)}$$

by

$$g_n^+(y) = P_{D_{a(v)}}\ \phi_L(x).$$

(We define $g_n^- : B_{\varepsilon_1}(x_n) \to K_{b(v)}$ in a similar way).



Thus for $x \in K_i$, x will lie in at least one ball in the cover and $g_n^+(x)$ will be defined. Thus we have replaced the discontinuous Poincaré map on $\bigcup_i K_i$ with an ambiguously defined version which is locally continuous, (for the right choice of image) ([2] p. 456).

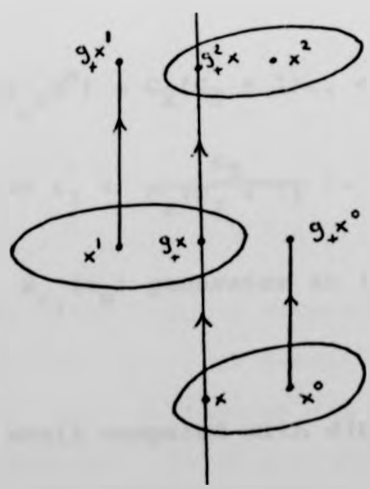
4.2 Construction of chains

We shall want to make use of the tracing property with ε_0-closeness of the tracing corresponding to ε_2-closeness of the chain. Providing we are judicious about our choice of (ε,T)-chains and trace sufficiently closely the ambiguous definition of g^+ will be unimportant (because we shall only need g^+ defined locally).

For $x \in \bigcup_{n=1}^{N} B_{\varepsilon_1}(x_n)$ we want to construct an (ε_2, T)-chain which takes values in the finite set $\{x_1, \ldots, x_N\}$. We proceed as follows:

Assume $x \in B_{\varepsilon_1}(x^o)$, where $x^o \in \{x_i\}$, then g^+ is defined so as to be continuous on $B_{\varepsilon_1}(x^o)$ and we may choose $x^1 \in \{x_n\}_{n=1}^{N}$ with $g^+(x) \in B_{\ 1}(x^1)$.

Proceeding inductively we construct $x^o, x^1, x^2, \ldots \in \{x_i\}$ with $d(x^n \ , (g^+)^n x) < \varepsilon_1$.



Let $g^+(x_n) = \phi_{t_n}(x_n)$ (since g_n^+ corresponds to flowing along the orbit of x).

Then:

$$d(\phi_{t^n}x^n, x^{n+1}) < d(\phi_{t^n}x^n, (g^+)^{n+1}x) + d((g^+)^{n+1}x, x^{n+1})$$

$$< (C_3 + 1)\varepsilon_1 < \varepsilon_2 \quad \text{(by Lemma 4)}, \ n \geq 0.$$

(provided we choose $\varepsilon_1 < \dfrac{\varepsilon_2}{C_3 + 1}$).

Repeating the argument with ϕ_{-t} we can construct sequences

$$x^0, x^{-1}, x^{-2}, \ldots \in \{x_i\} \text{ and } t^{-1}, t^{-2}, t^{-3}, \ldots \in \mathbb{R}^+$$

with

$$d(\phi_{-t^n}x^{n+1}, x^n) < (C_3 + 1)\varepsilon_1 \qquad (n < 0)$$

and so

$$d(x^{n+1}, \phi_{t^n}x^n) < C_2(C_3 + 1)\varepsilon_1 < \varepsilon_2 \quad \text{(by Lemma 3)}$$

(provided we choose $\varepsilon_1 < \dfrac{\varepsilon_2}{C_2(C_3 + 1)}$).

Thus $x \in \bigcup_{n=1}^{N} B_{\varepsilon_1}(x_n)$ generates an (ε_2, T)-chain.

<u>Note</u>:

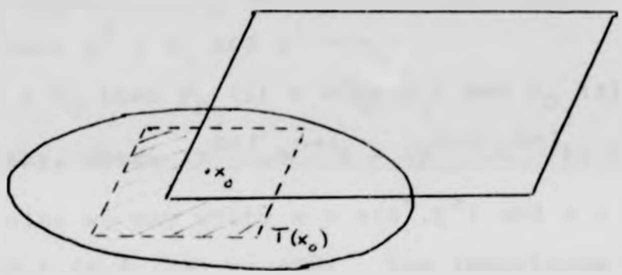
T is chosen small compared with $d(D_i, D_j)$, $(i \neq j)$.

4.3 <u>The Markov Partition</u>

We now start on the construction of the Markov partition (c.f. [5] §3.C, [2] §7).

Let $S = \{(x^n, t^n) : d(\phi_{t^n} x^n, x^{n+1}) < \varepsilon_2\}$ and

$$S_i = \{(x^n, t^n) \in S : x^0 = x_i\}.$$



For (x^n, t^n), $(y^n, s^n) \in S_i$ define

$$\langle y^n, s^n; x^n, t^n\rangle$$

$$= (\ldots y^{-1}, x_i, x^1, \ldots; \ldots s^{-1}, t^0, t^1 \ldots) \in S_i.$$

Given $(x^n, t^n) \in S$ the tracing property provides $x \in \Lambda \cap D_i$ which ε_0-traces (x^n, t^n). If ε_0 is small with respect to the expansive constant δ, say, (corresponding to some $\varepsilon << T$) then x is uniquely defined. Write $x = \theta(x^n, t^n)$ then $\theta\langle y^n, s^n; x^n, t^n\rangle = W^s_\varepsilon(x, T_i) \cap W^u_\varepsilon(y, T_i)$ [5].

If $\theta(s_i) = T_i$ then T_i is a rectangle.

<u>Remark</u>

Notice in particular that

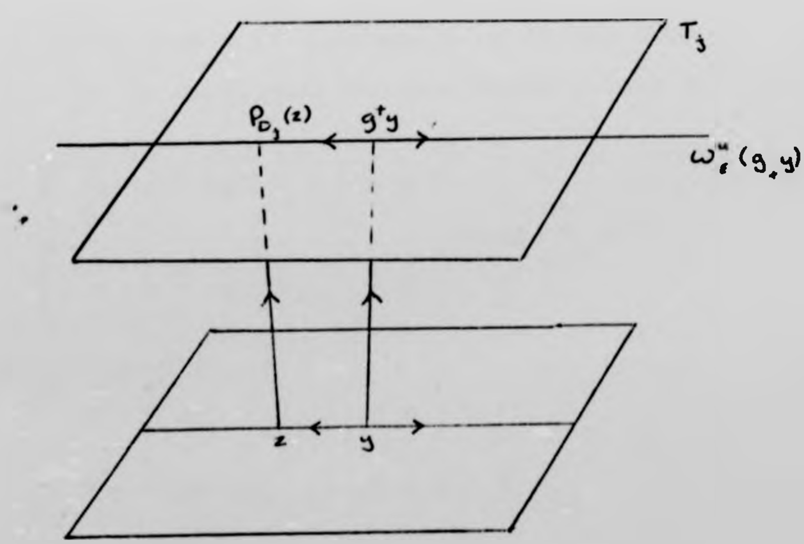
$$W^s_\varepsilon(x,T_i) = \theta\{(z^n,t^n) : z^n = x^n,\ n > 0\}$$

$$W^u_\varepsilon(x,T_i) = \theta\{(z^n,t^n) : z^n = x^n,\ n < 0\}.$$

If $z \in W^u_\varepsilon(y,D_i)$ and $y \in T_i$ and $g_+y \in T_j$ then write $y = \theta(y^n,t^n)$ where $y^0 = x_i$ and $y^1 = x_j$.

If $P_{D_j}(z) \in T_j$ then $P_{D_j}(z) \in W^u(g_+y,T_j)$ and $P_{D_j}(z) = \theta(z^{n+1},s^{n+1})$, say, where $(z^{n+1},s^{n+1}) = (y^{n+1},t^{n+1})$, $n < 0$.

In particular we may write $z = \theta(z^n_*,S^n)$ and $z \in T_i$ i.e. $z \in W^u(y,T_i)$ (c.f. [2] p. 458). The importance of this remark will be seen in the next section.



We know that $pr_{D_i} \langle\ ,\ \rangle : (W^s_\varepsilon(x) \cap \Lambda) \times (W^u_\varepsilon(x) \cap \Lambda) \to D_i$ is a homeomorphism onto its image. From this it is possible to show that the rectangles are proper i.e. $\overline{\text{int } T_i} = T_i$ ([2] p. 434).

Also $\text{int } T_i = pr_{D_i} \langle \text{int } W^u(x,T_i), \text{int } W^s(x,T_i)\rangle$ and $\text{int } T_i$ is also a rectangle (invariant under $pr_{D_i} \langle\ ,\ \rangle$) ([2], p. 432).

If we define

$$\partial^s R = \{x \in R : x \notin \text{int } W^u(x,R)\}$$

$$\partial^u R = \{x \in R : x \notin \text{int } W^s(x,R)\}$$

then $\partial R = \partial^s R \cup \partial^u R$, where $R \in \{T_i\}$ ([2], p. 432).

§5. SYMBOLIC DYNAMICS

We are now able to start on the construction of the suspension flow. Given a point $x \in \Lambda$, its orbit is determined by the list of transverse rectangles it intersects (future and past). Its position in the orbit is then given by the time since last passing through $\bigcup_i T_i$.

Let $H : \bigcup_i T_i \to \bigcup_i T_i$ be the Poincaré map.

5.1 Constructing the shift-space

Define a shift-space Σ by:

$$\Sigma = \{(T_{s(n)})_{n=-\infty}^{+\infty} : \forall \ell, k > 0, \bigcap_{j=-k}^{\ell} H^{-j}\text{int } T_{s(j)} \neq \emptyset\}$$

We claim that Σ is a subshift of finite type (c.f. [2], p. 437). It is sufficient to show there exists $N > 0$ s.t. if

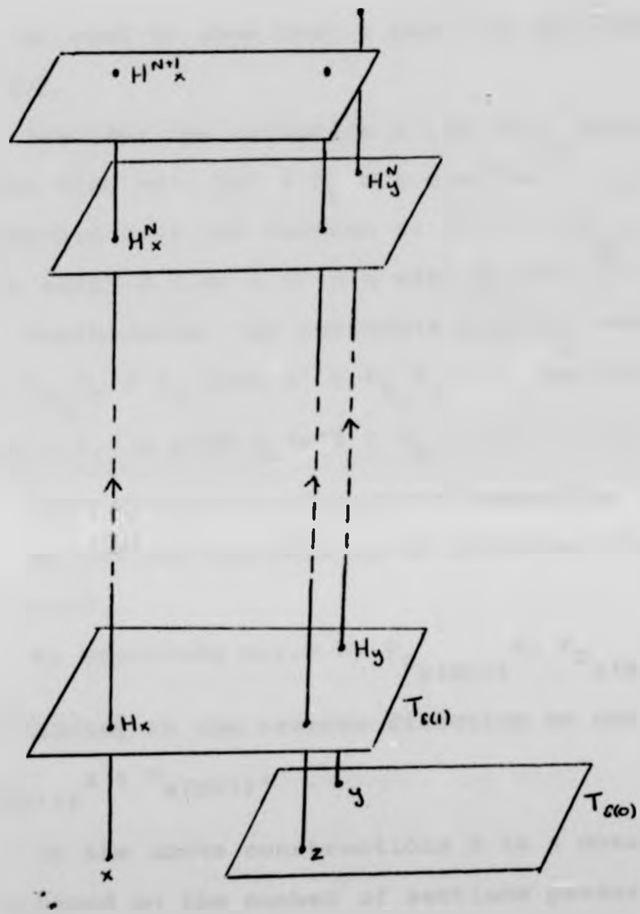
$$\left\{\begin{array}{l} \bigcap_{j=0}^{N} H^{-j} \text{ int } T_{s(j)} \neq \emptyset \\ \bigcap_{j=1}^{N+1} H^{-j} \text{ int } T_{s(j)} \neq \emptyset \end{array}\right. \quad \text{then } \bigcap_{j=0}^{N+1} H^{-j} \text{ int } T_{s(j)} \neq \emptyset$$

Assume

$$H^j x \in \text{int } T_{s(j)}, \quad 1 < j < N+1$$

$$H^j y \in \text{int } T_{s(j)}, \quad 0 < j < N$$

then $z = \langle x,y \rangle \in \bigcap_{j=0}^{N+1} H^{-j} \text{ int } T_{s(j)}$.



This needs elaboration:

(a) If $g^+y = H^t y \in T_{s(t)}$ then $P_{D_{s(t)}} z = P_{D_{s(t)}} \langle H^t x, H^t y\rangle \in T_{s(t)}$.

By a Remark in the preceeding section we have $z \in W^u(y, T_{s(0)})$, and in particular $z \in T_{s(0)}$ (c.f. [2], p. 438).

(b) We need to show that z meets no new rectangles (c.f. [2], p. 458).

Consider the situation $x', y' \in T_j$ then $z' = \rho_{D_j}\langle x',y'\rangle \in T_j$. Assume that $Hx', Hy' \in T_k$ but that $Hz' \in T_i$ ($\neq T_k$). By assumption (iv) (of Section 4) $x',y' \in P_{D_j}T_k \cap T_j$. (Thus there exist $0 < s' < t' < \alpha$ with $\phi_{s'}(z') \in T_i$; $\phi_{t'}(z') \in T_k$).

Furthermore, the rectangle property means that for $z' \in P_{D_j}T_i \cap T_j$ then $x' \in P_{D_j}T_i \cap T_j$ (so there exists $0 < s < t < \alpha$ with $\phi_s(x') \in T_k$, $\phi_t(x') \in T_i$).

This situation contradicts assumption (ii) of Section 3.

We can now use this as an inductive step to complete the proof.

By replacing x,y,z by $P_{D_{s(N+1)}}x$, $P_{D_{s(N+1)}}y$, $P_{D_{s(N+1)}}z$ and flowing in the reverse direction we can show $P_{D_{s(N+1)}}z \in T_{s(N+1)}$.

In the above constructions N is a constant (determined as a bound on the number of sections passed through on flowing from x to $g_+(x)$, etc).

Thus we have shown Σ to be a subshift of finite-type.

<u>Remark</u>

By a suitable refinement of the Markov partition (induced by a recoding of Σ) we may assume that Σ is given by a matrix A.

5.2 Invariance of boundaries

If we define a set $\Delta^s = \{\phi_t x : x \in \partial^s, 0 < t \leq 2L\}$, then we claim that $\phi_r \Delta^s \subseteq \Delta^s$ (for all $r > 0$) (c.f. [2], p. 439).

For $x \in T_i$, choose $j, x_n \in \text{int } T_i \cap H^{-1} \text{ int } T_j$ with $x_n \to x$. Then $A(i,j) = 1$ and $x \in T_i \cap H^{-1} T_j$. Thus $HW^u(x, T_i) \supseteq W^u(Hx, T_j)$. If $Hx \notin \partial^s$ then $W^u(Hx, T_j)$ is a neighbourhood of Hx in $W^s(Hx) \cap \Lambda$. Thus $W^u(x, T_i)$ is a neighbourhood of x in $W^s_\varepsilon(x) \cap \Lambda$. This means that $x \notin \partial^s$. This proves the claim. Similarly $\phi_{-r} \Delta^u \subseteq \Delta^u$ (for all $r > 0$).

5.3 The semi-conjugacy map

We define a map $\pi: \Sigma \to \bigcup_i T_i$ by $\pi(T_{s(n)}) = \bigcap_{n=-\infty}^{+\infty} H^{-n} \overline{\text{int } T_{s(n)}}$.

That this is non-empty follows from the construction of the shift-space. Furthermore the intersection consists of a single point by expansiveness. It also follows from the expansiveness of the flow that π is Lipschitz i.e. If $T_{s(n)} = T_{r(n)}$, $|n| < N$ then $d(\pi(T_{s(n)}), \pi(T_{r(n)})) < C\theta^N$ (Fixed $C > 0$, $0 < \theta < 1$) (c.f. [2], p. 435).

Because π is one-one on the dense set $R - \bigcup_{n=-\infty}^{\infty} H^{-n}(\partial^u \cup \partial^s)$ and $\pi(\Sigma)$ is compact we see that π is surjective.

To construct the suspended function let $r: \bigcup_i T_i \to \mathbb{R}^+$ be the return time for the flow i.e. $\phi_{r(x)} x = Hx$. Since the sections D_i and the flow are both C^1 we have r is Lipschitz on $T_i \cap H^{-1} T_j \subseteq D_i$.

Since $\pi:\Sigma \to \bigcup_i T_i$ is also Lipschitz, so is the composition $f = r\pi:\Sigma \to \mathbb{R}^+$.

We define the suspension space

$$\Sigma^f = \{(x,t) \in \Sigma \times \mathbb{R}^+ : 0 < t < f(x)\}$$

and identify $(x,f(x))$ and $(\sigma x,0)$. We define a flow on Σ^f by $\sigma_t^f(x,s) = (x,t+s)$ with identifications. We can extend π to $\pi:\Sigma^f \to \Lambda$ where $\pi(x,t) = \phi_t\pi(x,0)$.

Recalling the constructions we see $\pi\sigma_t^f = \phi_t\pi$ (c.f. [2], p. 436).

Remark

Since π is one-one on a dense set and ϕ_t is transitive it follows that σ_t^f is also transitive. In consequence $\sigma:\Sigma \to \Sigma$ is transitive.

5.4 The semi-conjugacy map is bounded-one

We shall now show π is bounded-one on Σ. If Σ is a shift on ℓ-symbols then we show $\pi^{-1}(x)$ contains at most ℓ^2 sequences (c.f. [6], p. 14). Choose N large that for $\{\underline{x}^i\} = \pi^{-1}(x)$ the strings $(x_0^i,\ldots,x_N^i)$ differ. With Card $\pi^{-1}(x) > \ell^2 + 1$ there exists $\underline{x}$, $\underline{y} \in \pi^{-1}(x)$ with $x_0 = y_0$ and $x_N = y_N$.

Choose $u \in \bigcap_{n=0}^{N} H^{-n} \operatorname{int} T(x_n)$ and $v \in \bigcap_{n=0}^{N} H^{-n} \operatorname{int} T(y_n)$ and since $T(x_0) = T(y_0)$ then $z = \rho_{D_0}\langle u,v\rangle \in \operatorname{int} T(x_0)$. Since $z \in W^s(u, T(x_0))$ then $H^n z \in \operatorname{int} T(x_n)$. Since the orbits of u and v lie close to that of z we may replace u,v,z by $H^N u$, $H^N v$ and $H^N z = \rho_{D_n}\langle H^N u, H^N v\rangle$. This gives that $H^n z \in \operatorname{int} T(y_n)$. Thus $x_n = y_n$ $0 < n < N$. This contradicts $\underline{x} \neq \underline{y}$.

<u>Remark</u>

Since $\pi : \Sigma^f \to \Lambda$ is, by extension, bounded-one we have $h(\phi) = h(\sigma^f)$ [7].

<u>Theorem</u> (Bowen) [2]

Let $\phi_t : \Lambda \to \Lambda$ be a C^1 Axiom A flow restricted to a basic set. There exists a flow $\sigma_t^f : \Sigma^f \to \Sigma^f$ where Σ is a transitive subshift of finite type and $f : \Sigma \to \mathbb{R}^+$ is Lipschitz.

Furthermore there exists a continuous, surjective, bounded-one map $\pi : \Sigma^f \to \Lambda$ such that $\phi_t \pi = \pi \sigma_t^f$.

5.5 <u>Asymptotic comparisons of closed orbits</u>

We want to show that asymptotically counting closed is the same for both flows.

Let $v_\phi(t)$ be the number of closed orbits for ϕ of length less than t. We want to strengthen the following result:

Proposition (Bowen) [4]

For an Axiom A flow $h(\phi) = \lim \frac{1}{t} \log v_\phi(t)$.

The number of orbits on $Z = \{x : \mathrm{Card}\ \pi^{-1}(x) = 1\}$ grows at a faster rate than on $\Lambda - Z$:

If $x \notin Z$ and $\phi_{t_0} x = x$ then $x \in \Delta^s \cup \Delta^u$.

Furthermore since Δ^s is closed, ϕ_t-invariant $(t > 0)$, and strictly contained in Λ then $h(\phi_t|_{\Delta^s}) < h(\phi) - \varepsilon$, say.

The number of closed orbits on Δ^s grows at a rate bounded by the entropy (Bowen and Walters) [8], i.e.

$$\overline{\lim}\ 1/t \log v_{\phi|_{\Delta^s}}(t) < h(\phi) - \varepsilon.$$

Similarly for $v_{\phi|_{\Delta^u}}(t)$.

Thus we conclude that $v_\phi(t) \sim v_{\phi|_Z}(t)$.

We can use a similar argument for $_\sigma f$ by replacing Z, Δ^s, Δ^u by $\pi^{-1}Z$, $\pi^{-1}\Delta^s$, $\pi^{-1}\Delta^u$ to show that $v_{\sigma^f}(t) \sim v_{\sigma^f|_{\pi^{-1}Z}}(t)$.

Conclusion: $v_\phi(t) \sim v_{\phi|_Z}(t) = v_{\sigma^f|_{\pi^{-1}Z}}(t) \sim v_{\sigma^f}(t)$ and asymptotically $v_\phi(t)$ and $v_{\sigma^f}(t)$ are the same (up to $O(e^{-\varepsilon t} v_\varphi(t))$.). (c.f. [2], p. 453).

§6. MANNING'S LEMMA

We now show explicitly the connection between closed orbits in Λ and Σ^f ([2] §5).

6.1 Preimages of the semi-conjugacy

Let $\tau \subseteq \Lambda$ be a chosen closed orbit of length $\lambda(\tau)$. Let $x \in \tau$ be a point on the orbit which does not lie on a rectangle.

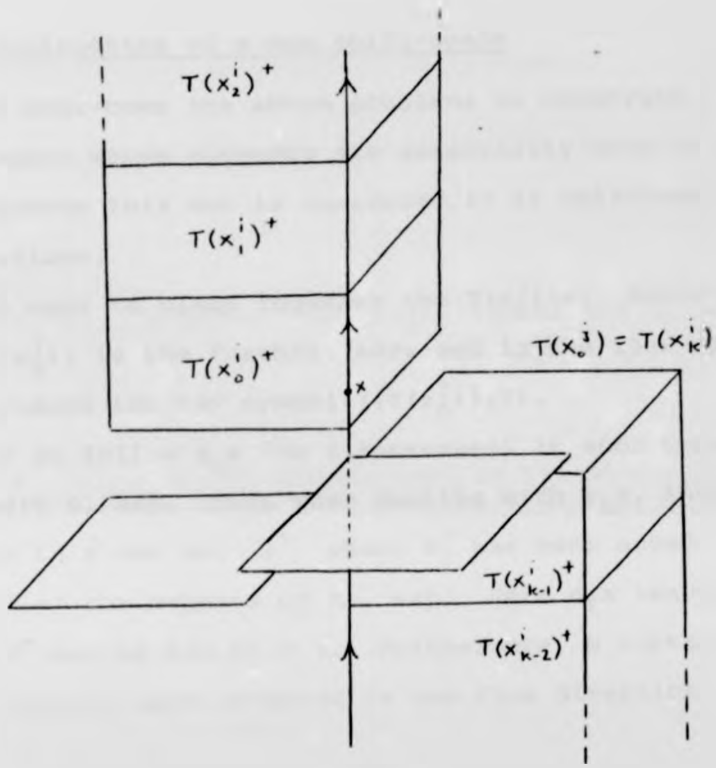
Let Card $\pi^{-1}(x) = N$ and $\pi^{-1}(x) = \{(x^i,t^i)\}$.

In some sense, x lies 'between' the rectangles $T(x_0^i)$ and $T(x_1^i)$ $(i = 1,\dots,N)$ and $x \in T(x_0^i)^+ = \{(z,t): z \in T(x_0^i),\ 0 < t < f(z)\}$.

We claim that the $\{T(x_0^i)\}$ are distinct. This is because if $x_0^i = x_0^j$ and $\sigma^n \underline{x}^i = \underline{x}^i$; $\sigma^m \underline{x}^j = \underline{x}^j$ then $x_{nm}^i = x_{nm}^j$. By a previous argument it follows that the intervening terms are identical and so $\underline{x}^i = \underline{x}^j$. This contradiction gives $x_0^i \neq x_0^j$.

What may happen is that $(\underline{x}^i,t^i)$ need not have period $\lambda(\tau)$, but period $m\lambda(\tau)$ instead $(1 < m < N)$.

To show how this occurs fix a choice of initial rectangle $T(x_0^i)$. Following the orbit of $\phi_t x$ gives a string of rectangles $T(x_0^i)$, $T(x_1^i)$, $T(x_2^i),\dots,T(x_k^i)$. When $\phi_t x$ returns to $\phi_{\lambda(\tau)} x = x$ we may find that $T(x_k^i) = T(x_0^j) \neq T(x_0^i)$.



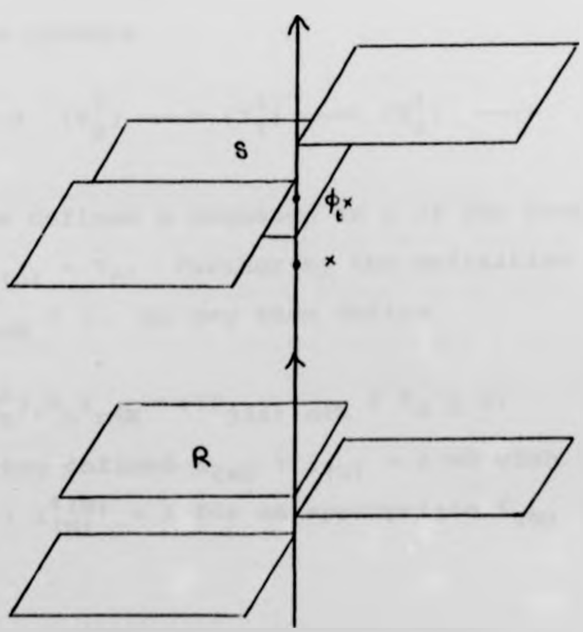
Since the $\{x_0^j\}$ are distinct the action of $\sigma^f_{\lambda(\tau)}$ permutes the $(\underline{x}^j, t^j)$ and consequently induces a permutation of the $\{T(x_0^j)\}$.

6.2 Construction of a new shift-space

To over-come the above problems we construct a new shift-space whose elements are essentially sets of rectangles. Thus because this set is unordered it is oblivious to such permutations.

We want to clump together the $T(x_0^j)$'s. Assume $T \in \{T(x_0^j)\}$ is the furthest advanced in the flow direction. We introduce the new symbol $(\{T(x_0^j)\},T)$.

If we follow $\phi_t x$ (as t increases) it soon transverses a rectangle S, say. Thus when dealing with $\phi_t x$, instead of x, it lies in a new set S^+ where S^+ has been added to $\{T(x_0^j)\}$ at the expense of R , say. Here $\phi_t x$ leaves R^+ to enter S^+ and so $A(R,S) = 1$. Furthermore, S instantly becomes the rectangle most advanced in the flow direction.



We represent this information as follows:

$$(\{T(x_0^j)\},T) \in \{(\{S^i\},S) : S \in \{S^i\},\ \exists x \in S,\ \varepsilon > 0$$
$$\text{s.t. } \phi_{[0,\varepsilon]}x \subseteq (S^i)^+ \text{ for each } i\}$$

and introduce a matrix for this (finite) set of symbols:

$$A_{(N)}(\{S^i\},S;\ \{T^i\},T) = 1$$

$\Longleftrightarrow$ $\{S^i\}$ and $\{T^i\}$ differ only in the change of one rectangle $S' \to T$, where $A(S',T) = 1$.

We can now use $A_{(N)}$ to construct a new shift space $\Sigma_{(N)}$. We now construct a map $\pi_{(N)}:\Sigma_{(N)} \to \Lambda$.

Let $(\{T_n^i\},T_n)_{n\in\mathbb{Z}} \in \Sigma_{(N)}$. Consider how T_0 evolves under the changes

$$\cdots \longrightarrow \{T_0^i\} \longrightarrow \{T_1^i\} \longrightarrow \{T_2^i\} \longrightarrow \cdots$$

This defines a sequence in Σ of the form $(T_{j(n)})_{n=-\infty}^{+\infty}$ where $T_{j(0)} = T_0$. Further by the definition of $A_{(N)}$ we have $(T_{j(n)})_{n\in\mathbb{Z}} \in \Sigma$. We may then define

$$\pi_{(N)}(\{T_n^i\},T_n)_{n\in\mathbb{Z}} = \pi(T_{j(n)})_{n\in\mathbb{Z}} \in T_0 \subseteq \Lambda.$$

Having defined $\pi_{(N)} : \Sigma_{(N)} \to \Lambda$ we wish to extend it to $\pi_{(N)} : \Sigma_{(N)}^{f(N)} \to \Lambda$ for an appropriate $f_{(N)} : \Sigma_{(N)} \to \mathbb{R}^+$.

Let t_0 be the flow time satisfying

$$\phi_{t_0}\pi_{(N)}(\{T_0^i\},T_0) = \pi_{(N)}(\{T_1^i\}\ ,\ T_1).$$

In fact t_0 is a flow time between points on rectangles T_0 and T_1, in Λ.

We define $f_{(N)} : \Sigma_{(N)} \to \mathbb{R}^+$ by $f_{(N)}(\{T_n^i\},T_n) = t_0$.
Because of the way in which $\pi_{(N)}, \Sigma_{(N)}, f_{(N)}$ have been constructed $\pi_{(N)}\sigma_t^{f(N)} = \phi_t\pi_{(N)}$ is always satisfied i.e. $\pi_{(N)}$ semi-conjugates the flows

$$\sigma_t^{f(N)} : \Sigma_{(N)}^{f(N)} \to \Sigma_{(N)}^{f(N)} \text{ and } \phi_t : \Lambda \to \Lambda.$$

The very incentive for this construction was that now $\pi_{(N)}^{-1}(x)$ consists of the single point $(\{T(x_0^j)\},T)$. Furthermore because $\{T(x_0^j)\}$ is insensitive to the permutations induced by $\sigma^f_{\lambda(\tau)}$ it follows that this single point has period $\lambda(\tau)$ under $\sigma^{f(N)}$.

6.3 Construction of other shift spaces

Unfortunately, as x varies then Card $\pi^{-1}(x)$ is prone to change. To accommodate these variations in the number of elements in $\pi^{-1}(x)$ we must introduce more shift-spaces. We first of all partition the finite set $\{T(x_0^j)\}$ e.g.

$$\{T(x_0^j)\} = \{T^i\} \cup \{T^j\} \cup \ldots \cup \{T^k\}$$

(where $\{T^i\} \cap \{T^j\} = \phi$, etc.)

Let t_0 be the flow time satisfying

$$\phi_{t_0}\pi_{(N)}(\{T_0^i\},T_0) = \pi_{(N)}(\{T_1^i\} \ , T_1).$$

In fact t_0 is a flow time between points on rectangles T_0 and T_1, in Λ.

We define $f_{(N)} : \Sigma_{(N)} \to \mathbb{R}^+$ by $f_{(N)}(\{T_n^i\},T_n) = t_0$.

Because of the way in which $\pi_{(N)}, \Sigma_{(N)}, f_{(N)}$ have been constructed $\pi_{(N)}\sigma_t^{f(N)} = \phi_t\pi_{(N)}$ is always satisfied i.e. $\pi_{(N)}$ semi-conjugates the flows

$$\sigma_t^{f(N)} : \Sigma_{(N)}^{f(N)} \to \Sigma_{(N)}^{f(N)} \text{ and } \phi_t : \Lambda \to \Lambda.$$

The very incentive for this construction was that now $\pi_{(N)}^{-1}(x)$ consists of the single point $(\{T(x_0^j)\},T)$. Furthermore because $\{T(x_0^j)\}$ is insensitive to the permutations induced by $\sigma^f_{\lambda(\tau)}$ it follows that this single point has period $\lambda(\tau)$ under $\sigma^{f(N)}$.

6.3 <u>Construction of other shift spaces</u>

Unfortunately, as x varies then Card $\pi^{-1}(x)$ is prone to change. To accommodate these variations in the number of elements in $\pi^{-1}(x)$ we must introduce more shift-spaces. We first of all partition the finite set $\{T(x_0^j)\}$ e.g.

$$\{T(x_0^j)\} = \{T^i\} \cup \{T^j\} \cup \ldots \cup \{T^k\}$$

(where $\{T^i\} \cap \{T^j\} = \phi$, etc.)

Next we form n-tuples whose terms are elements from the partition

i.e. we introduce $(\{T^i\}, \{T^j\},\ldots,\{T^k\};T)$

and again we emphasize the most advanced rectangle in the flow, T.

<u>Remark</u>

In the case of a 1-tuple we get $(\{T^i\};T)$ which is precisely the sort of symbol used in the previous sub-section.

To keep track of these partitions we introduce notation for indices:

Let $\underline{i} = (r_1,\ldots,r_n) \in \mathbb{Z}_+^n$ and write $\lambda(\underline{i}) = n$. Here $r_1 = \text{Card}\ \{T^i\}$; $r_2 = \text{Card}\ \{T^j\}$, etc.

By analogy with the case for $\Sigma_{(N)}$ we want to construct a shift space $\Sigma_{\underline{i}}$. We want sequences in $\Sigma_{\underline{i}}$ to take the form $(\{T_n^i\},\ldots,\{T_n^k\};T_n)$. Here, as usual, $T_n \in \{T_n^i\} \cup \ldots \cup \{T_n^k\}$ is the leading rectangle.

The admissibility condition for a sequence in $\Sigma_{\underline{i}}$ is that for

$$(\{T_m^i\},\ldots,\{T_m^k\};T_m) \longrightarrow (\{T_{m+1}^i\},\ldots,\{T_{m+1}^k\};T_{m+1})$$

all the n-tuples are the same, except one. Furthermore, in the entry that changes e.g. $\{T_m^j\} \longrightarrow \{T_{m+1}^j\}$ then exactly one rectangle $T \in \{T_m^j\}$ is replaced by a new rectangle R, say, satisfying $A(T,R) = 1$.

Remark

We see that these admissibility conditions are simply a generalisation of those for $\Sigma_{(N)}$.

By analogy with the case for $\underline{i} = (N)$ we may define $f_{\underline{i}}:\Sigma_{\underline{i}} \longrightarrow \mathbb{R}^+$ and $\pi_{\underline{i}}:\Sigma_{\underline{i}}^{f_{\underline{i}}} \longrightarrow \Lambda$ satisfying $\pi_{\underline{i}}\,\sigma_t^{f_{\underline{i}}} = \phi_t\pi_{\underline{i}}$.

If $(\{T_n^i\},\ldots,\{T_n^k\};T_n;t) \in \pi_{\underline{i}}^{-1}(x)$ then in order that this point has period $\lambda(\tau)$ we need that $(\{T_0^i\},\ldots,\{T_0^k\})$ coincides with the cycles in the permutation of $\{T(x_0^i)\}$ under $\sigma_{\lambda(\tau)}^f$.

6.4 The counting lemma

Assume $\underline{i} = (r_1,\ldots,r_n)$ and denote $\ell(\underline{i}) = n$ and $r = r_1 + \ldots + r_n$.

We define $N(t,\phi)$ to be the number of ϕ-closed orbits (primitive orbits) of length t in Λ. Similarly, let $N(t,\sigma^{f_{\underline{i}}})$ be the number of $\sigma^{f_{\underline{i}}}$-closed orbits in $\Sigma^{f_{\underline{i}}}$.

Theorem (Bowen, after Manning ([2], p. 449)).

$$N(t,\phi) = \sum_{\underline{i}} (-1)^{\ell(\underline{i})+1}\, N(t,\sigma^{f_{\underline{i}}})$$

We need only consider a particular $\tau \subseteq \Lambda$, and then only a single point x on this orbit. (This is because Card $\pi^{-1}(x)$ is constant on orbits).

Thus we only need to show

$$1 = \sum_{\underline{i}} (-1)^{\ell(\underline{i})+1} \text{ Card } \{\omega \in \pi_{\underline{i}}^{-1}(x) : \sigma_{\lambda(\tau)}^{f_{\underline{i}}}\omega = \omega\}.$$

We have already noticed that for ω to have period $\lambda(\tau)$ we need $(\{T_0^i\},\ldots,\{T_0^k\})$ to correspond to cycles in the $\sigma_{\lambda(\tau)}^f$ permutation of $\{T(x_0^i)\}$. If this is satisfied where $\underline{i}$ has $r < N$ then we may compare $\Sigma_{\underline{i}}^{f_{\underline{i}}}$ with $\Sigma_{\underline{i}^*}^{f_{\underline{i}^*}}$ where $\underline{i}^*$ is given by $\underline{i}^* = (r_1,r_2,\ldots,r_n,N-r)$. The periodic orbit in $\Sigma_{\underline{i}}^{f_{\underline{i}}}$ has a corresponding closed orbit in $\Sigma_{\underline{i}^*}^{f_{\underline{i}^*}}$. However since $\ell(\underline{i}^*) = \ell(\underline{i}) + 1$ their contributions to $N(\lambda(\tau),\phi)$ will cancel in the above theorem.

The only periodic orbit saved from this cancellation is the single orbit of period $\lambda(\tau)$ in $\pi_{(N)}^{-1}(x) \subseteq \Sigma_{(N)}^{f(N)}$. This is the closed orbit we considered in the preceding sub-section. This concludes the proof.

REFERENCES

[1] V.I. Arnold and A. Avez, *Ergodic problems of classical mechanics*, Benjamin, New York, 1968.

[2] R. Bowen, Symbolic dynamics for hyperbolic flows, Amer. J. Math, 95 (1973) 429-459.

[3] R. Bowen, Equidistribution of closed geodesics, Amer. J. Math, 94 (1972) 413-423.

[4] R. Bowen, Periodic orbits for hyperbolic flows, Amer. J. Math, 94 (1972) 1-30.

[5] R. Bowen, *Equilibrium states and the ergodic theory of Anosov diffeomorphisms*, S.L.N. 470, Springer-Verlag, Berlin, 1975.

[6] R. Bowen, *On Axiom A diffeomorphisms*, Amer. Math. Soc. Regional Conf. Proc., No. 35, 1978.

[7] R. Bowen, Entropy for group endomorphisms and homogeneous spaces, Trans. Amer. Math. Soc., 153 (1971) 401-414.

[8] R. Bowen and P. Walters, Expansive one-parameter flows, J. Differential Equations, 12 (1972) 180-193.

[9] J.E. Franke and J.F. Selgrade, Hyperbolicity and Chain Recurrence, J. Differential Equations, 26 (1977) 27-36.

[10] M. Hirsch, J. Palis, C. Pugh and M. Shub, Neighbourhoods of hyperbolic sets, Inventiones Math., 9 (1970) 121-134.

[11] M. Hirsch, C. Pugh and M. Shub, Invariant Manifolds, Bull. Amer. Math. Soc., 76 (1970) 1015-1019.

[12] C.C. Pugh and M. Shub, The Ω-stability theorem for flows, Inventiones Math., 11 (1970) 150-158.

[13] S. Smale, Differentiable dynamical systems, Bull. Amer. Math. Soc., 73 (1967) 747-817.